George Polya

# Matemáticas y razonamiento plausible

ESTRUCTURA Y FUNCION

EDITORIAL TECNOS, S. A. - MADRID

MATEMÁTICAS Y RAZONAMIENTO PLAUSIBLE es la obra principal de Polya. La línea central del libro es un estudio del razonamiento plausible, considerado por el autor como la lógica del descubrimiento matemático. Ha sido usual considerar que el razonamiento plausible o heurístico constituye la antítesis del razonamiento matemático, que debe ser demostrativo. La obra de Polya sostiene lo contrario, haciendo ver el importante e indispensable papel que lo plausible ha jugado siempre en la creación matemática.

Por otro lado, esta obra representa un notable esfuerzo en el campo de la fundamentación lógica de la matemática. El autor sostiene la tesis de que, en contra de lo que muchos han creído, la mayoría de los grandes descubrimientos matemáticos tienen su base en la experiencia. La importancia de la observación empírica y de sus secuelas: analogía e inducción, acercan la labor del matemático y la del naturalista. Poner esto de relieve en diferentes campos, desde la Teoría de los Números a la Geometría sólida, es el cometido del autor en estas páginas interesantísimas. Por si fuera poco, el autor nos ofrece una serie de *patrones* muy útiles en la investigación inductiva.

La primera parte del libro estudia el papel de la inducción, la analogía y la generalización en la invención matemática, utilizando siempre como método principal la intuición. La mayor parte de estos primeros capítulos tratan de mostrar en acción la importancia de la intuición en el descubrimiento científico; para ello, el autor dedica capítulos a la teoría de los números, la geometría sólida, problemas de óptica y de máximos y míni-

# Matemáticas y razonamiento plausible

# ESTRUCTURA Y FUNCION

*EL PORVENIR ACTUAL DE LA CIENCIA*

COLECCION DIRIGIDA POR

ENRIQUE TIERNO GALVAN

Núm. 1. *Quentin Gibson.*—LA LOGICA DE LA INVESTIGACION SOCIAL.—(2.ª edición.)
» 2. *M. F. Ashley Montagu.*—LA DIRECCION DEL DESARROLLO HUMANO
» 3. *Gustav Bergmann.*—FILOSOFIA DE LA CIENCIA
» 4. *Christian Bay.*—LA ESTRUCTURA DE LA LIBERTAD
» 5. *Ernest Nagel.*—LA LOGICA SIN METAFISICA
» 6. *David Hilbert y Wilhelm Ackermann.*—ELEMENTOS DE LOGICA TEORICA
» 7. *Vernon Van Dyke.*—CIENCIA POLITICA: UN ANALISIS FILOSOFICO
» 8. *Karl R. Popper.*—LA LOGICA DE LA INVESTIGACION CIENTIFICA
» 9. *R. M. Martin.*—VERDAD Y DENOTACION
» 10. *Ernest Gellner.*—PALABRAS Y COSAS
» 11. *C. D. Broad.*—EL PENSAMIENTO CIENTIFICO
» 12. *John Hospers.*—LA CONDUCTA HUMANA
» 13. *Arthur Pap.*—TEORIA ANALITICA DEL CONOCIMIENTO
» 14. *T. C. Schelling.*—LA ESTRATEGIA DEL CONFLICTO
» 15. *Richard B. Braithwaite.*—LA EXPLICACION CIENTIFICA
» 16. *Milic Capek.*—EL IMPACTO FILOSOFICO DE LA FISICA CONTEMPORANEA.
» 17. *Ros Ashby.*—PROYECTO PARA UN CEREBRO
» 18. *Hans Reichenbach.*—LA MODERNA FILOSOFIA DE LA CIENCIA

*EN PREPARACION:*

*D. M. Armstrong.*—LA PERCEPCION Y EL MUNDO FISICO
*G. L. S. Shackle.*—DECISION, ORDEN Y TIEMPO
*Curry y Feys.*—LOGICA COMBINATORIA
*Nagel-Newmann.*—LA PRUEBA DE GÖDEL
*Ernest Nagel.*—RAZON SOBERANA
*Max Black.*—MODELOS Y METAFORAS
*Ludwig Wittgenstein.*—"LOS LIBROS AZUL Y PARDO"
*Gödel.*—ENSAYOS FILOSOFICO-MATEMATICOS
*Jean Ladrière.*—LIMITACIONES INTERNAS DE LOS FORMALISMOS
*Henry Margenau.*—LA NATURALEZA DE LA REALIDAD FISICA: UNA FILOSOFIA DE LA FISICA MODERNA.

G. POLYA

# Matemáticas y razonamiento plausible

ESTRUCTURA Y FUNCION

*EDITORIAL TECNOS, S. A. - MADRID*

Los derechos para la versión castellana del libro
MATHEMATICS AND PLAUSIBLE REASONING,
publicado por *Princeton University Press*, de Princeton, New Jersey,
son propiedad de EDITORIAL TECNOS, S. A.

TRADUCCIÓN POR
JOSE LUIS ABELLAN

© EDITORIAL TECNOS, S. A., 1966
Calle O'Donnell, núm. 27. Teléfono 2 25 61 92. Madrid (9)
Número de registro: 5.741-64
Depósito legal: M.597-1966

SUCESORES DE RIVADENEYRA, S. A.—PASEO DE ONÉSIMO REDONDO, 26.—MADRID-8

# Tabla de materias

Páginas

PRÓLOGO ..... 13

INDICACIONES AL LECTOR ..... 21

PRIMERA PARTE

INDUCCION Y ANALOGIA EN MATEMATICAS

CAPÍTULO I. *Inducción* ..... 25

1. Experiencia y creencia.—2. Contactos que sugieren.—3. Contactos que apoyan.—4. La actitud inductiva.

Ejemplos y comentarios, 1-14. [12. Sí y No.—13. Experiencia y conducta.—14. El lógico, el matemático, el físico, el ingeniero.]

CAPÍTULO II. *Generalización, especialización, analogía* ..... 37

1. Generalización, especialización, analogía e inducción.—2. Generalización.—3. Especialización.—4. Analogía.—5. Generalización, especialización y analogía.—6. Descubrimiento por analogía.—7. Analogía e inducción.

Ejemplos y comentarios, 1-46. [Primera parte, 1-20; segunda parte, 21-46.] [1. La generalización correcta.—5. Un caso particular extremo.—7. Un caso particular guía.—10. Un caso particular representativo.—11. Un caso análogo.—18. Grandes analogías.—19. Analogías clarificadas.—20. Algunas citas.—21. La conjetura *E*.—44. Una objeción y una primera aproximación a la prueba.—45. Una segunda aproximación a la prueba.—46. Peligros de la analogía.]

CAPÍTULO III. *Inducción en geometría sólida* ..... 65

1. Poliedros.—2. Primeros contactos «de apoyo».—3. Más contactos «de apoyo».—4. Una prueba rigurosa.—5. Verificaciones y verificaciones.—6. Un caso muy diferente.—7. Analogía.—8. La división del espacio.—9. Modificando el problema.—10. Generalización, especialización, analogía.—11. Un problema análogo.—12. Una colección de problemas análogos.—13. Muchos problemas pueden ser más fáciles que uno solo.—14. Una conjetura.—15. Predicción y verificación.—16. Otra vez y mejor.—17. La inducción sugiere la deducción; el caso particular sugiere la prueba general.—18. Más conjeturas.

Páginas

Ejemplos y comentarios, 1-41. [21. Inducción: adaptación de la mente, adaptación del lenguaje.—31. La obra de Descartes sobre poliedros.—36. Angulos sólidos suplementarios y polígonos esféricos suplementarios.]

CAPÍTULO IV. *La inducción en la teoría de los números* ..... 95

1. Triángulos rectángulos en enteros.—2. Sumas de cuadrados.—3. Sobre la suma de cuatro cuadrados impares.—4. Examinando un ejemplo.—5. Tabulando las observaciones.—6. ¿Cuál es la regla?—7. Sobre la naturaleza del descubrimiento inductivo.—8. Sobre la naturaleza de la evidencia inductiva.

Ejemplos y comentarios, 1-26. [1. Notación.—26. Peligros de inducción.]

CAPÍTULO V. *Ejemplos variados de inducción* ..... 115

1. Desarrollo.—2. Aproximaciones.—3. Límites.—4. Intentando refutarlo.—5. Intentando probarlo.—6. El papel de la fase inductiva.

Ejemplos y comentarios, 1-18. [15. Explicar las regularidades observadas.—16. Clasificar los hechos observados.—18. ¿Cuál es la diferencia?]

CAPÍTULO VI. *Un enunciado más general* ..... 133

1. Euler.—2. La memoria de Euler.—3. Transición a un punto de vista más general.—4. Bosquejo esquemático de la memoria de Euler.

Ejemplos y comentarios, 1-25. [1. Funciones generantes.—7. Un problema combinatorio en geometría plana.—10. Suma de cuadrados.—19. Otra fórmula periódica.—20. Otra ley muy extraordinaria de los números referente a la suma de sus divisores.—24. Cómo Euler erró un descubrimiento.—25. Una generalización del teorema de Euler sobre $\sigma(n)$.]

CAPÍTULO VII. *La inducción matemática* ..... 152

1. La fase inductiva.—2. La fase demostrativa.—3. Examinando transiciones.—4. La técnica de la inducción matemática.

Ejemplos y comentarios, 1-18. [12. Probar más puede traer menos problemas.—14. Equilibre su teorema.—15. Perspectiva.—17. ¿Son iguales cualesquiera $n$ números?]

CAPÍTULO VIII. *Máximos y mínimos* ..... 169

1. Patrones.—2. Ejemplo.—3. El patrón de la línea de nivel tangente.—4. Ejemplos.—5. El patrón de variación parcial.—6. El teorema de los medios aritméticos y geométricos y sus primeras consecuencias.

Ejemplos y comentarios, 1-63. [Primera parte, 1-32; segunda parte, 33-63.] [1. Las distancias máximas y mínimas en geometría plana.—2. Las distancias máximas y mínimas en geometría sólida.—3. Líneas de nivel en un plano.—4. Superficies de nivel en el espacio.—11. El principio de la línea de paso a nivel.—22. El principio de variación parcial.—23. Existencia del extremo.—24. Una modificación del patrón de variación parcial: Un proceso infinito.—25. Otra modificación del patrón de variación parcial: Un proceso finito.—26. Confrontación gráfica.—33. Polígonos y poliedros. Area y perímetro. Volumen y superficie.—34. Prisma recto con base cuadrada.—35. Cilindro recto.—36. Prisma recto general.—

Páginas

37. Pirámide doble recta con base cuadrada.—38. Cono recto doble.—39. Pirámide doble recta general.—43. Aplicando la geometría al álgebra.—45. Aplicando el álgebra a la geometría.—51. Pirámide recta con base cuadrada.—52. Cono recto.—53. Pirámide recta general.—55. La caja sin tapa.—56. La artesa.—57. Un fragmento.—62. Un problema de correos.—63. Un problema de Kepler.]

CAPÍTULO IX. *Matemática física*..... 168

1. Interpretación óptica.—2. Interpretación mecánica.—3. Reinterpretación.—4. El descubrimiento por Bernouille de la braquistocrona.—5. El descubrimiento por Arquímedes del cálculo integral.

Ejemplos y comentarios, 1-38. [3. Triángulo con perímetro mínimo inscrito en un triángulo dado.—9. Centro de tráfico con cuatro puntos en el espacio.—10. Centro de tráfico con cuatro puntos en el plano.—11. Red de tráfico con cuatro puntos.—12. Desdoblar y estirar.—13. Billares.—14. Exploración geofísica.—23. Las líneas más cortas en una superficie poliédrica.—24. Las líneas más cortas (geodésicas) en una superficie curvada.—26. Una construcción con papel doblado.—27. La suerte está echada.—28. El Diluvio.—29. No tan profundo como un pozo.—30. Un caso extremo útil.—32. El cálculo de variaciones.—33. Del equilibrio de las secciones transversales al equilibrio de los sólidos.—38. El método de Arquímedes, retrospectivamente.]

CAPÍTULO X. *El problema isoperimétrico*..... 227

1. Las razones inductivas de Descartes.—2. Razones latentes.—3. Razones físicas.—4. Las razones inductivas de lord Rayleigh.—5. Derivando consecuencias.—6. Verificando consecuencias.—7. Muy cerca.—8. Tres formas del teorema isoperimétrico.—9. Aplicaciones y problemas.

Ejemplos y comentarios, 1-43. [Primera parte, 1-15; segunda parte, 16-43.] [1. Mirando atrás.—2. ¿Podríamos derivar parte del resultado de modo diferente?—3. Volver a formular con más detalle.—7. ¿Podemos utilizar el método para algún otro problema?—8. Formas perspicaces del teorema isoperimétrico.—16. La cuerda y el palo.—21. Dos palos y dos cuerdas.—25. El problema de Dido en geometría sólida.—27. Bisectores de una región del plano.—34. Bisectores de una superficie cerrada.—40. Una figura de muchas perfecciones.—41. Un caso análogo.—42.—Los sólidos regulares.—43. Razones inductivas.]

CAPÍTULO XI. *Más clases de razones plausibles*..... 253

1. Conjeturas y conjeturas.—2. Juzgando por un caso afín.—3. Juzgando por el caso general.—4. Prefiriendo la conjetura más sencilla.—5. Concepción del mundo.—6. Inagotable.—7. Asunciones heurísticas usuales.

Ejemplos y comentarios, 1-23. [16. El caso general.—19. Ninguna idea es realmente mala.—20. Algunas asunciones heurísticas.—21. Optimismo compensatorio.—23. El cálculo numérico y el ingeniero.]

OBSERVACIÓN..... 277

Páginas

SEGUNDA PARTE

## PATRONES DE INFERENCIA PLAUSIBLE

CAPÍTULO XII. *Algunos patrones notables* ..... 281

1. Verificación de una consecuencia.—2. Verificación sucesiva de varias consecuencias.—3. Verificación de una consecuencia improbable.—4. Inferencia por analogía.—5. Profundización de la analogía.—6. Inferencia analógica matizada.

Ejemplos y comentarios, 1-14. [14. Conclusión inductiva de esfuerzos sin fruto.]

CAPÍTULO XIII. *Otros patrones y primeros lazos* ..... 299

1. Examinando una consecuencia.—2. Examinando un fundamento posible.—3. Examinando una conjetura conflictiva.—4. Términos lógicos.—5. Lazos lógicos entre patrones de inferencia plausible.—6. Inferencia matizada.—7. Una tabla.—8. Combinación de patrones sencillos.—9. Sobre la inferencia por analogía.—10. Inferencia cualificada.—11. Sobre verificaciones sucesivas.—12. Sobre conjeturas rivales.—13. Sobre la prueba judicial.

Ejemplos y comentarios, 1-20. [Primera parte, 1-10; segunda parte, 11-21.] [9. Sobre la investigación inductiva en matemáticas y en las ciencias físicas.—10. Formulaciones generales tentativas.—11. Más personal, más complejo.—12. Tenemos una línea recta que une dos puntos dados.—13. Hay una línea recta con una dirección dada por un punto dado. Dibujar una paralela.—14. El caso más obvio puede ser el único posible.—15. Formando el modelo. El poder de las palabras.—16. Esto es demasiado improbable para ser una mera coincidencia.—17. Perfeccionando la analogía.—18. Una nueva conjetura.—19. Otra nueva conjetura.—20. ¿Qué es típico.]

CAPÍTULO XIV. *El azar, conjetura siempre presente* ..... 342

1. Fenómenos azarosos de masas.—2. El concepto de probabilidad.—3. Usando el saco y las bolas.—4. El cálculo de probabilidad. Hipótesis estadística.—5. Predicción correcta de frecuencias.—6. Explicación de fenómenos.—7. Juzgando hipótesis estadísticas.—8. Eligiendo entre hipótesis estadísticas.—9. Juzgando conjeturas no estadísticas.—10. Juzgando conjeturas matemáticas.

Ejemplos y comentarios, 1-33. [Primera parte, 1-18; segunda parte, 19-33.] [19. Sobre el concepto de probabilidad.—20. Cómo no debe ser interpretado el concepto de frecuencia de la probabilidad.—24. La probabilidad y la solución de problemas.—25. Regular e irregular.—26. Las reglas fundamentales del cálculo de probabilidad.—27. Independencia.—30. Permutaciones a partir de la probabilidad.—31. Combinaciones a partir de la probabilidad.—32. La elección de una conjetura estadística rival: un ejemplo.—33. La elección de una conjetura estadística rival: observaciones generales.]

Páginas

CAPÍTULO XV. *El cálculo de probabilidad y la lógica del razonamiento plausible*.. 406

1. Reglas de razonamiento plausible.—2. Un aspecto del razonamiento demostrativo.—3. Un aspecto similar del razonamiento plausible.—4. Un aspecto del cálculo de probabilidad. Dificultades.—5. Un aspecto del cálculo de probabilidad. Un intento.—6. Examinando una consecuencia.—7. Examinando un fundamento posible.—8. Examen de conjeturas en conflicto.—9. Examinando varias consecuencias en sucesión.—10. Sobre la evidencia circunstancial.

Ejemplos y comentarios, 1-9. [4. Probabilidad y credibilidad.—5. Verosimilitud y credibilidad.—6. Un intento de Laplace para enlazar inducción con probabilidad.—7. ¿Por qué no cuantitativo?—8. ¿Credibilidades infinitesimales?—9. Reglas de admisibilidad.]

CAPÍTULO XVI. *El razonamiento plausible en la invención y en la enseñanza*.. 445

1. Objeto del presente capítulo.—2. La historia de un pequeño descubrimiento.—3. El proceso de solución.—4. «Deus ex machina».—5. Justificación heurística.—6. La historia de otro descubrimiento.—7. Algunas indicaciones típicas.—8. La inducción en la invención.—9. Algunas palabras al profesor.

Ejemplos y comentarios, 1-13. [1. Para el profesor: algunos tipos de problemas.—7. «Qui nimium probat, nihil probat».—8. Proximidad y credibilidad.—9. Cálculo numérico y razonamiento plausible.—13. Demostración formal y razonamiento plausible.]

SOLUCIONES A PROBLEMAS ..... 477

BIBLIOGRAFÍA ..... 579

# *Prólogo*

*Este libro tiene diversos objetivos, estrechamente relacionados entre sí. En primer lugar, trata de servir a profesores y estudiantes de matemáticas en algo importante, aunque generalmente desdeñado. Por otra parte, el libro es un ensayo filosófico y también una continuación de trabajos anteriores que requiere, a su vez, ser continuada. Tocaré todos estos puntos, uno tras otro.*

*1. Estrictamente hablando, todos nuestros conocimientos, aparte de las matemáticas y de la lógica demostrativa (que es, en realidad, una rama de aquéllas), consisten en conjeturas. Hay, naturalmente, conjeturas y conjeturas. Unas, que merecen todo respeto y confianza, como las expresadas en ciertas leyes generales de la física, y otras, que no merecen ni confianza ni respeto, algunas de las cuales nos llenan de ira al leerlas en los periódicos. Y, entre unas y otras, se dan toda clase de conjeturas, presentimientos e intuiciones.*

*Aseguramos nuestro conocimiento matemático mediante el* razonamiento demostrativo, *pero apoyamos nuestras conjeturas por medio del* razonamiento plausible. *Una prueba matemática es razonamiento demostrativo, pero la evidencia inductiva del físico, la evidencia circunstancial del abogado, la evidencia documental del historiador y la evidencia estadística del economista pertenecen al razonamiento plausible.*

*Hay grandes diferencias entre las dos clases de razonamiento.*

*El razonamiento demostrativo es seguro, definitivo, y está más allá de toda controversia. El razonamiento plausible es azaroso, discutible y provisional. Aquél penetra las ciencias naturales tanto como la matemática, pero es, en sí mismo, igual que esta última, incapaz de producir un conocimiento esencialmente nuevo sobre el mundo en torno. Para aprender algo nuevo sobre el mundo necesitamos el razonamiento plausible, que es la única clase de razonamiento que utilizamos en nuestra vida cotidiana. El razonamiento demostrativo tiene modelos rígidos, codificados y aclarados por la lógica (formal o demostrativa), que es la teoría del razonamiento demostrativo. Los modelos del razonamiento plausible son fluidos y no hay teoría*

*de este razonamiento que pueda ser comparada a la lógica demostrativa en claridad o que tenga un consenso comparable.*

*2. Hay otro punto concerniente a estas dos clases de razonamiento, que merece nuestra atención. Todos sabemos que las matemáticas ofrecen una excelente oportunidad de aprender el razonamiento demostrativo, pero yo sostengo también que no hay materia en los programas usuales de las escuelas que ofrezca una oportunidad semejante de aprender el razonamiento plausible. Me dirijo a todos los estudiantes interesados en matemáticas de todos los grados y les digo: «Aprendamos a probar, desde luego, pero* aprendamos también a intuir».

*Esto suena un poco paradójico y debo hacer hincapié en algunos puntos para evitar posibles malos entendidos.*

*Las matemáticas son consideradas como una ciencia demostrativa. Sin embargo, éste es sólo uno de sus aspectos. La obra matemática se nos presenta, una vez terminada, como puramente demostrativa, consistente en pruebas solamente. No obstante, esta ciencia se asemeja en su desarrollo al de cualquier otro conocimiento humano. Hay que intuir un teorema matemático antes de probarlo, así como la idea de la prueba antes de llevar a cabo los detalles. Hay que combinar observaciones, seguir analogías y probar una y otra vez. El resultado de la labor demostrativa del matemático es el razonamiento demostrativo, la prueba, pero ésta a su vez es descubierta mediante el razonamiento plausible, mediante la intuición. Si el aprendizaje de las matemáticas refleja en algún grado la invención de esta ciencia debe haber en él un lugar para la intuición, para la inferencia pausible.*

*Hemos dicho que hay dos clases de razonamiento: demostrativo y plausible. Permítaseme observar que no se contradicen entre sí; por el contrario, se completan uno al otro. En el razonamiento estricto lo principal es distinguir una prueba de una intuición, una demostración válida de un intento sin validez. En el razonamiento plausible lo importante es distinguir entre intuiciones, unas más y otras menos razonables. Si dirigimos nuestra atención a estas distinciones, ambas pueden hacerse más claras.*

*Un estudiante seriamente interesado en matemáticas, que pretenda dedicar a ellas su vida, debe aprender el razonamiento demostrativo; él es su profesión y el signo distintivo de su ciencia. Sin embargo, para obtener un éxito real debe también aprender el razonamiento plausible; de él dependerá su labor creadora. El estudiante aficionado tomará también el gusto al razonamiento demostrativo: quizá tenga poca necesidad de usarlo directamente, pero con él adquirirá un término con el que comparar la supuesta evidencia de todas clases con que se enfrentará en la vida moderna. Sin embargo, en todos sus esfuerzos necesitará del razona-*

*miento plausible. En todo caso, un estudiante de matemáticas ambicioso intentará aprender las dos clases de razonamiento, demostrativo y plausible, cualesquiera que puedan ser sus intereses ulteriores.*

*3. No creo que haya un método a toda prueba para aprender a intuir. De cualquier modo, si tal método existe yo no lo conozco y, desde luego, no pretendo ofrecerlo en las siguientes páginas. El uso eficiente del razonamiento plausible es una habilidad práctica, aprendida, como el resto de ellas, mediante la imitación y el uso. Intentaré hacer lo que pueda por el lector ansioso de aprender el razonamiento plausible, pero todo lo que puedo ofrecerle son sólo ejemplos que imitar, y darle una oportunidad de practicarlo.*

*En lo que sigue examinaré con frecuencia grandes y pequeños descubrimientos matemáticos. No puedo contar la historia real de cómo ocurrió el descubrimiento porque eso nadie lo sabe. No obstante, intentaré reflejarlo lo mejor posible, haciendo hincapié en los motivos que subyacen en el descubrimiento, en las inferencias plausibles que condujeron a él y, en resumen, en todo lo que merezca ser imitado. Naturalmente, trataré de impresionar al lector; ése es mi deber como profesor y autor. Sin embargo, seré totalmente honrado con él en la cuestión que realmente importa: intentaré impresionarle sólo con cosas que parezcan legítimas y me sirvan de ayuda.*

*Cada capítulo irá seguido de ejemplos y comentarios. Estos tratarán de puntos demasiado técnicos o sutiles para el texto del capítulo, o que resulten un poco marginales al hilo principal del argumento. Algunos de los ejercicios dan al lector la oportunidad de volver a considerar detalles sólo esbozados en el texto. Sin embargo, la mayoría de los ejercicios están pensados para que el lector saque las conclusiones por sí mismo. Antes de enfrentarse con un problema propuesto al final del capítulo el lector leerá cuidadosamente las partes más importantes de dicho capítulo y echará también una ojeada a los problemas inmediatos; en unas u otros puede estar contenida una clave del problema. Con objeto de que tales claves sean del mayor beneficio para la instrucción del lector se ha cuidado mucho no sólo del contenido y la forma de los problemas propuestos, sino también de su* disposición. *En efecto, se nos ha ido mucho más tiempo en la ordenación de estos problemas de lo que podría imaginarse.*

*Para alcanzar un amplio círculo de lectores he intentado ilustrar cada cuestión importante con un ejemplo tan elemental como me ha sido posible. No obstante, en varios casos me vi obligado a poner ejemplos no demasiado elementales que apoyasen la cuestión con suficiente fuerza. Me pareció que debía presentar también ejemplos de interés histórico, ejemplos de belleza matemática y otros que ilustrasen el paralelismo con los procedimientos de otras ciencias, o de la vida diaria.*

*Añadiré que la forma final de muchas de las historias contadas resultó de una especie de experimento psicológico. Discutí el tema con los alumnos en diferentes clases, interrumpiendo mi exposición frecuentemente con preguntas como la siguiente: «Bien, ¿qué harían ustedes en tal situación?». Varios pasajes incorporados en el texto han sido sugeridos por sus respuestas, o he modificado mi versión original de alguna manera por la reacción de mi auditorio.*

*En resumen, he tratado de utilizar toda mi experiencia en la investigación y en la enseñanza para dar al lector una adecuada oportunidad de imitar inteligentemente y de hacer las cosas por sí mismo.*

*4. Los ejemplos de razonamiento plausible recogidos en este libro pueden tener otro uso como iluminadores de un problema filosófico muy discutido: el problema de la inducción. La cuestión crucial es: ¿Existen reglas para la inducción? Algunos filósofos lo afirman, pero la mayor parte de los científicos cree que no existen tales reglas. Con objeto de discutir esta cuestión provechosamente la plantearemos de diferente modo. En su tratamiento confiaremos poco en los verbalismos tradicionales, o en formalismos a ultranza, manteniéndonos en estrecho contacto con la práctica de los científicos. Ahora bien, observemos que el razonamiento inductivo es un caso particular del razonamiento plausible, y observemos también (lo que los autores modernos olvidan casi siempre, pero que algunos escritores antiguos, como Euler y Laplace, percibieron con claridad) que el papel de la evidencia inductiva en la investigación matemática es semejante a su papel en la investigación física. Así podremos advertir la posibilidad de obtener alguna información sobre el razonamiento inductivo mediante la observación y comparación de ejemplos de razonamiento plausible en las cuestiones matemáticas. De esta manera se nos abre la puerta para* investigar la inducción inductivamente.

*Cuando un biólogo pretende investigar algún problema general, digamos de genética, es muy importante que elija algunas especies particulares de plantas o animales que se presten bien al estudio experimental de su problema. El químico que quiere investigar un problema general, como es, por ejemplo, la velocidad de las reacciones químicas, debe elegir algunas sustancias especiales en las que poder hacer convenientemente los experimentos relevantes a este problema. La elección de material experimental apropiado es de gran importancia en la investigación inductiva de cualquier problema.*

*Me parece que las matemáticas son, en diversos aspectos, el material experimental más apropiado para el estudio del razonamiento inductivo. Este estudio comprende experimentos psicológicos de cierta clase: sentimos que nuestra confianza en una conjetura se basa en varias clases de evidencia.*

*Gracias a su inherente sencillez y a su claridad, las cuestiones matemáticas se presentan a esta especie de experimento psicológico mucho mejor que los temas de cualquier otro campo. En las siguientes páginas el lector encontrará una amplia oportunidad para convencerse de esto por sí mismo.*

*Considero más filosófico examinar la idea general del razonamiento plausible en vez del caso particular del razonamiento inductivo. Los ejemplos recogidos en este libro conducen a un aspecto definido y claramente satisfactorio del razonamiento plausible. Sin embargo, no deseo inculcar mis puntos de vista al lector. En efecto, ni siquiera los establezco en la primera parte del libro; quiero que los ejemplos hablen por sí mismos. Los primeros cuatro capítulos de la segunda parte están dedicados a un examen general más explícito del razonamiento plausible. Allí enuncio formalmente los patrones de inferencia plausible sugeridos por los ejemplos precedentes, e intento sistematizar dichos patrones y observar algunas de sus relaciones entre ellos mismos, y con la idea de probabilidad.*

*No sé si el contenido de estos cuatro capítulos merece ser llamado filosofía. Si lo es, se trata de una modesta fisosofía, más dedicada a la comprensión de ejemplos concretos y a la conducta concreta de la gente que a la exposición de generalidades. Naturalmente, no conozco cuál será la opinión final que merezca mi posición. Sin embargo, tengo bastante confianza en que mis ejemplos puedan ser útiles a todos los estudiantes de inducción o de razonamiento plausible, que no tengan prejuicios, y deseen formar sus puntos de vista en estrecho contacto con los hechos observables.*

5. *Esta obra sobre* Matemáticas y razonamiento plausible *se divide en dos partes:* Inducción y analogía en matemáticas, *la primera, y* Patrones de inferencia plausible, *la segunda. En la primera parte se encuentra casi todo el núcleo matemático de la obra, así como los «datos» para la investigación inductiva de la inducción, que se realiza en la segunda parte. No he dado un índice, pensando que esto contribuiría a hacer la terminología más rígida de lo que es deseable en esta clase de obras. Creo que la tabla de materias proporcionará una guía satisfactoria del libro.*

*La presente obra es una continuación de mi primer libro* How to Solve It *. *El lector interesado en el tema deberá leer ambos, sin importar el orden en que lo haga. Este texto está ordenado de manera que puede leerse independientemente del primer trabajo. En efecto, hay aquí sólo unas cuantas referencias directas al primer libro, que no es necesario considerar en una primera lectura. Sin embargo, las referencias indirectas a aquel libro están en casi cada página de éste e incluso en cada frase en algunas*

* *Cómo resolverlo;* no está traducido al español.—*N. del T.*

*páginas. De hecho, en esta obra hay numerosos ejercicios y algunas ilustraciones más avanzadas para las que no hubo espacio en el primer libro, debido a su tamaño y a su carácter elemental.*

*También nos referimos aquí a una colección de problemas de* Analysis, *obra de G. Szegö y del autor de ésta (ver bibliografía). Los problemas de esta colección se han ordenado cuidadosamente en series, de forma que se apoyan mutuamente, se proporcionan claves unos a otros, cubriendo entre todos una determinada materia y dando al lector la oportunidad de practicar varios pasos importantes en la resolución de problemas. En el tratamiento de éstos el presente texto sigue el método de presentación iniciado en la primera obra, y este hecho no deja de tener importancia.*

*Dos capítulos de la segunda parte del presente libro tratan de la teoría de la probabilidad. El primero de estos capítulos tiene cierta conexión con una exposición elemental del cálculo de probabilidades escrita por el autor hace varios años (ver bibliografía). Los puntos de vista que subyacen en este tema, así como el punto de partida, son los mismos, pero en los demás hay poca relación entre estas obras.*

*Algunas de las posiciones mantenidas en este libro han sido expresadas antes en mis ensayos citados en la bibliografía. En el texto han sido incorporados extensos pasajes de mis ensayos números 4, 6, 8, 9 y 10. Agradezco a los editores de la* American Mathematical Monthly, Etudes de Philosophie des Sciences en Hommage à Ferdinand Gonseth, *y* Proceedings of the International Congress of Mathematicians 1950, *quienes amablemente permitieron que fuesen reimpresos estos pasajes.*

*Muchas partes de este libro han sido expuestas en mis conferencias y algunas de ellas varias veces. En ciertas partes y en algunos aspectos he conservado el tono de la exposición oral. No creo que tal tono sea aconsejable en la presentación impresa de las matemáticas en general, pero en el presente caso puede ser apropiado o al menos excusable.*

*6. El contenido del último capítulo de este libro, que trata de investigación y enseñanza, enlaza más explícitamente con la primera obra del autor e indica una posible conclusión.*

*El uso eficiente del razonamiento plausible juega un papel fundamental en la resolución de problemas. Hemos intentado ilustrar este papel con muchos ejemplos, pero hay otros aspectos de esta cuestión que necesitan una ilustración semejante.*

*Muchos puntos tratados aquí necesitan posteriores trabajos. Sería conveniente confrontar mi posición con la de otros autores, examinar más profundamente los ejemplos históricos, investigar, hasta donde sea posible, mis puntos de vista sobre la investigación y la enseñanza*

*con los métodos de la psicología experimental* *, *etc. Quedan por hacerse varias de estas tareas, aunque algunas de ellas pueden resultar ingratas.*

*El presente no es un libro de texto. No obstante, espero que con el tiempo influirá en la presentación usual de los libros de texto y en la elección de sus problemas. La tarea de re-escribir los libros de texto de temas usuales, al modo sugerido en estas líneas, creo que contaría con el agradecimiento de todos.*

7. *Deseo expresar mi gratitud a la* Princeton University Press *por la impresión cuidadadosa del texto y especialmente a Mr. Herbert S. Bailey, Jr., director de la misma, por su comprensiva ayuda en varias cuestiones. También estoy en deuda con Mrs. Priscilla Feigen, por haber pasado a máquina el original, y con el doctor Julius G. Baron, por su amable ayuda al leer las pruebas.*

GEORGE POLYA

Stanford University
Mayo, 1953

---

* En esta dirección se han emprendido trabajos exploratorios en el Departamento de Psicología de la Universidad de Stanford dentro de un programa dirigido por E. R. Hilgard, bajo el patrocinio de O. N. R.

# Indicaciones al lector

La sección 2 del capítulo VII se cita como sección 2 en el mismo, pero se cita como sección 7.2 en los demás. La subsección (3) de la sección 5 del capítulo XIV se cita como sección 5 (3) en el mismo, pero se cita como sección 14.5 (3) en los restantes. Nos referimos al ejemplo 26 del capítulo XIV como ejemplo 26 en dicho capítulo, pero como ejemplo 14.26 en todos los demás capítulos.

Para leer las partes sustanciales del texto bastará algún conocimiento de álgebra elemental y geometría. Para casi el total del texto y la mayoría de los ejemplos y comentarios será suficiente un conocimiento completo de álgebra elemental y geometría, y algún conocimiento de geometría analítica y cálculo, incluidos límites y series infinitas. Sin embargo, se da por supuesto en el lector un conocimiento más avanzado en unas pocas observaciones incidentales del texto, en algunos problemas propuestos y en varios comentarios. Cuando se supone un conocimiento superior se hace, generalmente, alguna advertencia.

El lector más preparado, que salte aquellas partes que le parecen demasiado elementales, puede perder más que el lector menos preparado que pase por alto las partes que le resulten demasiado complejas.

A menudo se omiten sin advertirlo algunos detalles de demostraciones (no muy difíciles). Un lector debidamente preparado para esta eventualidad no necesita estropear sus buenos hábitos críticos.

La mayoría de los problemas propuestos para solución son muy fáciles, pero unos pocos son bastante difíciles. Las indicaciones que pueden facilitar su solución van encerradas entre corchetes: [ ]. Los problemas cercanos pueden dar también indicaciones. Se dará especial atención a las líneas introductorias que van delante de los ejemplos de algunos capítulos, o delante de la primera parte, o segunda parte, de tales ejemplos.

Las soluciones son a veces muy breves: dan por supuesto que el lector ha intentado resolver con ahínco el problema por sus propios medios antes de mirar a la solución impresa.

Un lector que se esfuerza seriamente en resolver un problema puede

obtener provecho por ello aunque no logre resolverlo. Por ejemplo, puede mirar la solución, intentar aislar la que le parece la idea clave, poner el libro a un lado y entonces intentar encontrar la solución.

En algunos puntos este libro es pródigo en figuras o en dar pequeños pasos intermedios de una derivación. La pretensión es hacer visible la *evolución* de una figura o unas fórmulas; véanse, por ejemplo, las figuras 16.1-16.5. Sin embargo, ningún libro puede tener bastantes figuras o fórmulas. Un lector puede desear leer un pasaje «por encima» o más cuidadosamente. Si desea leerlo así, tendrá papel y lápiz a mano, preparándose para escribir o trazar cualquier fórmula o figura dada o solamente indicada por el texto. Haciéndolo así, tiene más posibilidades de ver la evolución de la figura o fórmula, de comprender cómo los varios detalles contribuyen al resultado final y de recordar el conjunto.

# PRIMERA PARTE

# INDUCCION Y ANALOGIA EN MATEMATICAS

# Inducción

*Sin duda, ha de parecer bastante paradójico atribuir gran importancia a la observación en la parte de las matemáticas conocida como matemática pura, pues la opinion corriente cree que la observación debe restringirse a los objtos físicos. En cuanto referimos los números sólo al intelecto puro no comprendemos bien cómo la observación y los cuasi-experimentos pueden ser útiles en la investigación sobre la naturaleza de los números. Sin embargo, de hecho, como demostraré aquí con buenas razones, las propiedades de los números conocidos hoy han sido en su mayor parte descubiertas por observación, y mucho tiempos antes de que su verdad haya sido confirmada por rigugurosas demostraciones. Hay, por otro lado, muchas propiedades de los números con las que estamos familiarizados, sin que seamos todavía capaces de probarlas; sólo la observación nos ha conducido a su conocimiento. Por lo que vemos en la teoría de los números, aun siendo todavía muy imperfecta, podemos colocar nuestras más altas esperanzas en la observación; ella nos conduce continuamente a nuevas propiedades que pueden ser probadas posteriormente. Debemos distinguir cuidadosamente de la verdad el conocimiento que sólo se apoya en observaciones y no ha sido aún probado; se trata de un conocimiento obtenido por inducción, como usualmente decimos. Y, puesto que nosotros hemos visto casos en que la mera inducción conduce al error, debemos tener sumo cuidado en no aceptar como verdaderas las propiedades de los números que han sido descubiertas por observación y que se apoyan sólo sobre la inducción. En realidad, nosotros usaremos tales descubrimientos como una oportunidad para investigar más detenidamente las propiedades descubiertas y probarlas o refutarlas; en cualquiera de ambos casos aprenderemos algo útil.*—EULER *.

## 1. EXPERIENCIA Y CREENCIA

La experiencia modifica las creencias humanas. Nosotros aprendemos de la experiencia, o, mejor dicho, debíamos aprender de ella. Hacer el mejor uso posible de la experiencia es una de las grandes empresas humanas y trabajar por ella es la vocación de los científicos.

* EULER, *Opera Omnia*, ser. 1, vol. 2, pág. 459. Specimen de usu observationum in mathesi pura.

Un científico digno de este nombre tratará de extraer de una experiencia determinada las conclusiones más correctas y acumular las experiencias más útiles para establecer la mejor línea de investigación respecto a una cuestión dada. El procedimiento del científico para tratar con la experiencia se suele llamar *inducción*. Ejemplos de procedimiento inductivo especialmente claros los encontramos en la investigación matemática; en la próxima sección presentaremos uno muy sencillo.

## 2. Contactos que sugieren

La inducción empieza frecuentemente con alguna observación. Un naturalista observa la vida de los pájaros; un cristalógrafo, las formas de los cristales. Un matemático, interesado en la teoría de los números, observa las propiedades de los enteros 1, 2, 3, 4, 5, ...

Si usted desea observar la vida de los pájaros con alguna oportunidad de obtener interesantes resultados deberá de alguna manera familiarizarse con ellos, interesarse en ellos; quizá hasta lleguen a gustarle los pájaros. De manera semejante, si usted desea observar los números deberá interesarse y familiarizarse con ellos. Usted distinguirá entre números pares y nones; conocerá los cuadrados 1, 4, 9, 16, 25, ... y los primos 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, ... (Es mejor dejar aparte el 1 como «unidad» y no clasificarlo como primo.) Incluso con tan modesto conocimiento usted puede ser capaz de observar algo interesante.

Por casualidad usted puede tropezarse con estas relaciones:

$$3 + 7 = 10, \quad 3 + 17 = 20, \quad 13 + 17 = 30,$$

y observar alguna semejanza entre ellas. Le chocará que los números 3, 7, 13 y 17 son números primos impares. La suma de dos primos impares es necesariamente un número par; en efecto, 10, 20 y 30 son pares. ¿Qué pasa con los *otros* números pares? ¿Se comportan también de manera semejante? El primer número par que es una suma de dos primos nones es, sin duda,

$$6 = 3 + 3.$$

Si miramos más allá de 6 encontraremos que

$$\begin{aligned} 8 &= 3 + 5 \\ 10 &= 3 + 7 = 5 + 5 \\ 12 &= 5 + 7 \\ 14 &= 3 + 11 = 7 + 7 \\ 16 &= 3 + 13 = 5 + 11. \end{aligned}$$

¿Seguirá así siempre? De cualquier modo, los casos particulares observados sugieren un juicio general: *Cualquier número par mayor de 4 es la suma de dos primos impares.* Pero teniendo en cuenta las excepciones, 2 y 4, que no pueden dividirse en la suma de dos primos impares, es preferible este otro juicio un poco más complicado: *Cualquier número par que no es primo, ni cuadrado de primo, es la suma de dos primos impares.*

Hemos llegado así con ello a formular una *conjetura*, que ha sido encontrada por *inducción*. Esto es, sugerida por observación, indicada por ejemplos particulares.

Estas indicaciones son más bien endebles; de hecho, tenemos un fundamento muy débil para creer en nuestra conjetura. Sin embargo, encontraremos algún consuelo si sabemos que el matemático que descubrió esta conjetura hace más de doscientos años, Goldbach, no tenía fundamento mucho más fuerte para ella.

¿Es verdadera la conjetura de Goldbach? Nadie puede hoy contestar a esta pregunta. A pesar del gran esfuerzo hecho por algunos grandes matemáticos, la conjetura de Goldbach es hoy, como en los días de Euler, una de esas «muchas propiedades de los números con que estamos bien familiarizados, aunque seamos incapaces de probarlas» o de refutarlas.

Ahora, retrocedamos y tratemos de reconstruir los pasos dados en el razonamiento precedente como es típico del procedimiento inductivo.

Primero, *observamos alguna semejanza.* Nos dimos cuenta que 3, 7, 13 y 17 son primos, que 10, 20 y 30 son pares, y que las tres ecuaciones $3 + 7 = 10$, $3 + 17 = 20$, $13 + 17 = 30$ son *análogas* entre sí.

El segundo paso fue una *generalización.* De los ejemplos 3, 7, 13 y 17 pasamos a todos los primos impares; de 10, 20 y 30, a todos los números pares, y establecimos después una posible relación general:

$$\text{número par} = \text{primo} + \text{primo}.$$

De esta manera llegamos a un juicio general y claramente formulado, que, sin embargo, es meramente conjetural o *tentativo*. Es decir, el juicio no ha sido probado por ningún medio y no puede tener pretensión de certeza; es sólo un intento de alcanzar la verdad.

Esta conjetura tiene, a pesar de todo, algunos *sugestivos puntos de contacto* con la experiencia, con «los hechos» o con «la realidad». Es cierta para el caso particular de los pares 10, 20 y 30, y también para 6, 8, 12, 14, 16.

Con estas observaciones bosquejamos *grosso modo* un primer estadio de proceso inductivo.

## 3. Contactos que apoyan

Usted no pondrá mucha confianza en una conjetura que no ha sido probada, aunque haya sido propuesta por una gran autoridad, incluso aunque haya sido propuesta por usted mismo. Usted intentará demostrarla o refutarla; sin duda, la *probará*.

La conjetura de Goldbach será probada si examinamos algunos nuevos números pares y decidimos si son o no la suma de dos primos impares. Veamos, por ejemplo, el número 60, y realicemos un «cuasi experimento», como dice Euler. El número 60 es par, pero ¿es también la suma de dos primos? ¿Es cierto que

$$60 = 3 + \text{primo?}$$

No; 57 no es primo. ¿Es

$$60 = 5 + \text{primo?}$$

La contestación es, otra vez: «No»; 55 no es primo. Si seguimos por este camino, la conjetura será destruida. Sin embargo, el próximo ensayo da resultado:

$$60 = 7 + 53,$$

y 53 es un número primo. La conjetura ha sido comprobada en un caso más.

El resultado contrario habría destruido la conjetura de Goldbach de una vez por todas. Si ensayamos con todos los números primos bajo un número par dado, como 60, y nunca llegamos a poder descomponerlo en la suma de dos primos, la conjetura habrá sido destruida irrevocablemente. Por el contrario, si verificamos la conjetura en el caso del número

par 60, no hemos llegado con ello a ninguna conclusión definitiva. Indudablemente, no se demuestra un teorema por una sola verificación, lo que no obsta para interpretarla como un *signo favorable*. Por lo que se refiere a la conjetura, tal verificación le da *más crédito*, aunque, por supuesto, pertenece a la decisión personal el peso que se ha de dar a dicho signo.

Volvamos, por un momento, al número 60. Después de haber ensayado los primos 3, 5, y 7 podemos hacerlo ahora con los primos que quedan bajo 30. (Como es obvio, resulta innecesario ir más allá de 30, que es igual a 60/2, puesto que uno de los dos primos, cuya suma será 60, tiene que ser menor que 30.) Podemos, pues, descomponer 60 en las siguientes sumas de dos primos:

$$60 = 7 + 53 = 13 + 47 = 17 + 43 = 19 + 41 = 23 + 37 = 29 + 31.$$

Ahora podemos proceder sistemáticamente y examinar los números pares uno tras otro, como hemos hecho con el 60. Podemos *tabular* los resultados así:

$$\begin{aligned}
6 &= 3 + 3 \\
8 &= 3 + 5 \\
10 &= 3 + 7 = 5 + 5 \\
12 &= 5 + 7 \\
14 &= 3 + 11 = 7 + 7 \\
16 &= 3 + 13 = 5 + 11 \\
18 &= 5 + 13 = 7 + 11 \\
20 &= 3 + 17 = 7 + 13 \\
22 &= 3 + 19 = 5 + 17 = 11 + 11 \\
24 &= 5 + 19 = 7 + 17 = 11 + 13 \\
26 &= 3 + 23 = 7 + 19 = 13 + 13 \\
28 &= 5 + 23 = 11 + 17 \\
30 &= 7 + 23 = 11 + 19 = 13 + 17.
\end{aligned}$$

La conjetura queda verificada en todos los casos que hemos establecido aquí. Cada verificación que alarga la tabla fortalece la conjetura, le da más crédito, la hace más plausible. Naturalmente, no existe un número de verificaciones que demuestre la conjetura.

Ahora examinemos, comparemos y combinemos las observaciones recogidas y busquemos alguna clave que se oculte tras ella. En nuestro

caso es muy difícil descubrir alguna clave esencial en la tabla. Aun así, si la examinamos podemos darnos cuenta claramente de su significado.

La tabla nos muestra la frecuencia con que los números pares incluidos en ella pueden ser representados como una suma de dos primos (6 sólo una vez, 30 tres veces). El número de tales representaciones del número par $2n$ parece «incrementar irregularmente» con $n$. La conjetura de Goldbach expresa la esperanza de que el número de representaciones nunca caerá por debajo de 0, por mucho que la tabla se extienda.

De entre los casos que hemos examinado podemos distinguir dos grupos: el de los que preceden a la formulación de la conjetura y el de los que la siguen. Los primeros sugieren la conjetura, los segundos la apoyan. Ambos casos nos proporcionan alguna clase de contacto entre la conjetura y «los hechos». La tabla no distingue entre los puntos de contacto «que sugieren» y los «que apoyan».

Ahora volvamos al razonamiento precedente y tratemos de verlo según los rasgos típicos del proceso inductivo.

Una vez concebida una conjetura intentaremos descubrir si es verdadera o falsa. Nuestra conjetura era un juicio general sugerido por ciertos ejemplos particulares en que la habíamos encontrado verdadera. Luego examinamos otros ejemplos particulares y, puesto que la conjetura resultó verdadera en todos los casos examinados, nuestra confianza se incrementó.

Me parece que en todo lo anterior no hemos hecho más que lo que cualquier persona razonable haría. Y obrando así hemos llegado a un principio: *Un juicio general y conjetural adquiere más crédito si es verificado en un nuevo caso particular.*

¿Es éste el principio subyacente en todo proceso de inducción?

## 4. La actitud inductiva

En nuestra vida personal nos aferramos con frecuencia a ilusiones. En otras palabras: no nos atrevemos a examinar ciertas creencias, que podrían contradecirse con la experiencia, por temor a destruir nuestro equilibrio emocional. Pueden darse circunstancias en que no resulte insensato aferrarse a las ilusiones, pero en la ciencia es necesaria una actitud muy diferente: la *actitud inductiva*. Esta actitud reclama adaptar nuestras creencias y experiencias tan eficazmente como sea posible. Ello requiere una cierta preferencia por las cuestiones de hecho. Y requiere también saber ascender de las observaciones a las generalizaciones y descender de las generalizaciones más altas a las más concre-

tas observaciones. Sin duda, hemos de aprender a decir «puede ser» y «quizá» en miles de formas diferentes. Además, se requieren otras muchas cosas, pero principalmente estas tres:

Primero, estar dispuestos a revisar cualquiera de nuestras creencias.

Segundo, ser capaces de cambiar una creencia cuando existe una razón compulsiva para ello.

Tercero, no cambiar las creencias frívolamente sin que haya alguna buena razón.

Estos puntos parecen bastante triviales. Sin embargo, deben ir acompañados de cualidades poco frecuentes con arreglo a las cuales el científico debe vivir.

El primer punto necesita «coraje intelectual». Se necesita coraje para revisar las propias creencias. Galileo, cambiando los prejuicios de sus contemporáneos y la autoridad de Aristóteles, es un gran ejemplo de coraje intelectual.

El segundo necesita de «honestidad intelectual». Adherirse a una conjetura que ha sido claramente contradicha por la experiencia sólo porque es *mi* conjetura sería deshonesto.

El tercer punto necesita de «sabia contención». Cambiar una creencia sin serio examen, sólo por la dirección de la moda, pongamos como ejemplo, sería insensato. Aun así, no tenemos tiempo ni espacio para examinar seriamente todas nuestras creencias. Por tanto, lo cuerdo será abandonar nuestros problemas y nuestras dudas activas sobre las creencias hasta que tengamos una oportunidad razonable para aclararlas. «No creamos nada, pero preguntémonos sólo sobre aquello que merezca la pena preguntarse.»

Coraje intelectual, honestidad intelectual y sabia contención son las cualidades morales del científico.

## EJEMPLOS Y COMENTARIOS

**1.** Averiguar la norma con arreglo a la cual han sido elegidos los términos de la secuencia siguiente:

$$11, 31, 41, 61, 71, 101, 131, \ldots$$

**2.** Consideremos esta tabla:

$$\begin{aligned} 1 &= 0 + 1 \\ 2 + 3 + 4 &= 1 + 8 \\ 5 + 6 + 7 + 8 + 9 &= 8 + 27 \\ 10 + 11 + 12 + 13 + 14 + 15 + 16 &= 27 + 64. \end{aligned}$$

Averiguar la ley sugerida por estos ejemplos, expresarla en una adecuada notación matemática y demostrarla.

**3.** Observar los valores de las sumas sucesivas

$$1, \quad 1+3, \quad 1+3+5, \quad 1+3+5+7, \ldots$$

¿Hay aquí una regla sencilla?

**4.** Observar los valores de las sumas consecutivas

$$1, \quad 1+8, \quad 1+8+27, \quad 1+8+27+64, \ldots$$

¿Hay aquí una regla simple?

**5.** Los tres lados de un triángulo tienen de longitud $l$, $m$ y $n$, respectivamente. Las cifras $l$, $m$ y $n$ son enteros positivos, $l \leqq m \leqq n$. Encontremos el número de triángulos diferentes de la clase descrita por un $n$ dado. [Tomemos $n = 1, 2, 3, 4, 5, \ldots$] Encontremos una ley general que regule la subordinación del número de triángulos respecto de $n$.

**6.** Las primeras tres cifras de la secuencia 5, 15, 25, ... (números terminados en 5), son divisibles por 5. ¿Son también las cifras siguientes divisibles por 5?

Las primeras tres cifras de la secuencia 3, 13, 23, ... (números que terminan en 3), son números primos. ¿Son también las cifras siguientes números primos?

**7.** Por cómputo formal encontramos

$$(1 + 1!x + 2!x^2 + 3!x^3 + 4!x^4 + 5!x^5 + 6!x^6 + \ldots)^{-1}$$
$$= 1 - x - x^2 - 3x^3 - 13x^4 - 71x^5 - 461x^6 \ldots$$

Esto sugiere dos conjeturas sobre los siguientes coeficientes de las series de valores de la derecha: (1) todos son negativos; (2) todos son primos. ¿Son estas dos conjeturas igualmente dignas de crédito?

**8.** Serie

$$\left(1 - \frac{x}{1} + \frac{x^2}{2} - \frac{x^3}{3} + \ldots\right)^{-1} = A_0 + \frac{A_1 x}{1!} + \frac{A_2 x^2}{2!} + \ldots,$$

donde tenemos que para

$$n = 0\ 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9$$
$$A_n = 1\ 1\ 1\ 2\ 4\ 14\ 38\ 216\ 600\ 6240.$$

Establecer una conjetura.

**9.** El gran matemático francés Fermat consideró la secuencia

$$5, 17, 257, 65537, \ldots,$$

cuya cifra general es $2^{2^n} + 1$. Observó que los primeros cuatro números (dados aquí), correspondientes a $n = 1, 2, 3$ y 4, son primos, y conjeturó que los números siguientes serían también primos. Aunque él no lo demostró, se sintió tan seguro de su conjetura que desafió a Wallis y otros matemáticos ingleses a demostrarlo. Sin embargo, Euler encontró que la próxima cifra, $2^{32} + 1$, correspondiente a $n = 5$, no es un primo: ya que es divisible por 641 *. Veamos el pasaje de Euler al frente de este capítulo: «...nosotros hemos visto casos en que la mera inducción conduce a error».

**10.** Al verificar la conjetura de Goldbach para $2n = 60$ hemos ensayado sucesivamente los primos $p$ bajo $n = 30$. Sin embargo, podríamos haber también ensayado los primos $p'$ entre $n = 30$ y $2n = 60$. ¿Qué procedimiento es probablemente más ventajoso para la $n$ mayor?

**11.** En un diccionario nosotros encontraremos para la explicación de palabras como «inducción», «experimento» y «observación» frases parecidas a las siguientes:

«Inducción es inferir una ley general de ejemplos particulares, o una producción de hechos para demostrar un juicio general»;

«Experimento es un procedimiento para comprobar hipótesis»;

«Observación es una correcta apreciación y anotación de fenómenos, tal como ocurre en la Naturaleza respecto a la causa y el efecto o a sus relaciones mutuas».

¿Son estas descripciones aplicables al ejemplo discutido por nosotros en las secciones 2 y 3?

**12.** *Sí y No.* El matemático, como el naturalista, al comprobar alguna secuencia de una ley conjetural por medio de una nueva observa-

---

* EULER, *Opera Omnia*, ser. 1, vol. 2, págs. 1-5. HARDY y WRIGHT, *The Theory of Numbers*, págs. 14-15.

ción, dirige una pregunta a la Naturaleza: «Sospecho que esta ley es cierta. ¿Es cierta?». Si la consecuencia queda claramente rechazada, la ley no puede ser cierta. Si, por el contrario, la consecuencia es claramente verificada, hay indicio de que la ley puede ser cierta. La Naturaleza puede contestar Sí o No, pero sólo susurra una contestación mientras aniquila la otra; su Sí es provisional, su No es definitivo.

**13.** *Experiencia y conducta.* La experiencia modifica la conducta y las creencias humanas. Estas dos cosas no son independientes entre sí. La conducta es a menudo un resultado de las creencias; éstas son conducta potencial. Aun cuando usted puede ver la conducta de sus compañeros, usted no puede ver sus creencias. La conducta es más fácil de observar que las creencias. Todo el mundo sabe que «niño quemado, al fuego teme», que expresa justamente lo que decimos: la experiencia modifica la conducta humana.

Sí, y la conducta animal también.

En mi vecindad hay un perro que ladra y salta sobre todo el mundo sin que le provoquen. Pero he encontrado un medio de protegerme fácilmente. Si me paro e intento recoger una piedra el perro se aleja aullando. Todos los perros no se conducen de la misma manera, y es fácil adivinar qué clase de experiencia condujo a este perro a ese comportamiento.

El oso del Zoológico «pide comida». Es decir, cuando hay un espectador cercano se coloca en una ridícula postura que frecuentemente incita al espectador a lanzarle un terrón de azúcar en la jaula. Los osos que no están en cautividad probablemente no adoptan nunca tan absurdas posturas, y es fácil imaginar qué clase de experiencia condujo al oso del Zoo a iniciar tal comportamiento.

Una investigación completa de la inducción incluiría, sin duda, el estudio del comportamiento animal.

**14.** *El lógico, el matemático, el físico, el ingeniero.* «Mira ese matemático —dice el lógico—. El observa que los noventa y nueve primeros son números inferiores a 100 e infiere de aquí, por lo que él llama inducción, que todos los números son menores a 100.»

«Un físico cree —dice el matemático— que 60 es divisible por todos los números. El observa que 60 es divisible por 1, 2, 3, 4, 5 y 6; después examina algunos casos más, tales 10, 20 y 30, tomados al azar, como él dice. Y puesto que 60 es divisible también por éstos considera ya como suficiente la evidencia experimental.»

«Sí; pues mira al ingeniero —dice el físico—. Un ingeniero sospecha que todos los números impares son primos. "De cualquier manera, el 1

podemos considerarlo como un número primo", arguye. Luego vienen 3, 5 y 7, que son, indudablemente, primos. Después, 9; un caso de torpeza: no parece ser primo. Aun 11 y 13 son, indudablemente, primos. "Y retrocediendo al 9 —él dice—debemos concluir que es un error experimental".»

Es obvio que la inducción puede conducir a error. Sin embargo, hay que observar, ya que las posibilidades de error parecen tan abrumadoras, que la inducción algunas veces conduce a la verdad. ¿Empezaremos con el estudio de los casos evidentes, en que la inducción fracasa, o con aquellos en que tiene éxito? El estudio de las piedras preciosas es increíblemente más atractivo que el de los cantos vulgares y, por si fuera poco, son las piedras preciosas más que los cantos quienes conducen a los mineralogistas a la maravillosa ciencia de la cristalografía.

# Generalización, especialización, analogía

*Y yo estimo las analogías más que nada, son mis guías más dignas de confianza. Ellas conocen todos los secretos de la Naturaleza y debían ser menos descuidadas en geometría.*—KEPLER.

## 1. Generalización, especialización, analogía e inducción

Volvamos otra vez al ejemplo de razonamiento inductivo que hemos examinado con algún detalle (secciones 1.2, 1.3). Empezamos por observar la *analogía* de las tres relaciones

$$3 + 7 = 10, \quad 3 + 17 = 20, \quad 13 + 17 = 30;$$

después *generalizamos* ascendiendo desde 3, 7, 13 y 17 a todos los primos; desde 10, 20 y 30 a todos los números pares, y luego *especializamos* de nuevo, bajando a probar números particulares, tales como 6, 8 ó 60.

El primer ejemplo es extremadamente simple e ilustra de modo correcto el papel de la generalización, la especialización y la analogía en el razonamiento inductivo. Sin embargo, examinaremos ilustraciones más vivas, y antes de esto discutiremos la generalización, la especialización y la analogía, estas grandes fuentes de descubrimientos, por sí mismas.

## 2. Generalización

Es el paso de la consideración de una serie determinada de objetos a la de una serie mayor que contiene a la primera. Por ejemplo, generalizamos cuando pasamos de la consideración de triángulos a la de polígonos con arbitrario número de lados. Generalizamos también cuando pasamos del estudio de las funciones trigonométricas de un ángulo agudo a las funciones trigonométricas de un ángulo indeterminado.

Se puede observar que en estos dos ejemplos la generalización se ha efectuado de dos formas típicamente diferentes. En el primer ejemplo,

al pasar de triángulos a polígonos con $n$ lados, remplazamos una constante por una variable, el entero fijo 3 por el entero arbitrario $n$ (determinado sólo por la desigualdad $n \geqq 3$). En el segundo ejemplo, al pasar de ángulos agudos a ángulos arbitrarios $\alpha$, eliminamos una restricción, esto es, la restricción de $0° < \alpha < 90°$.

A menudo generalizamos al pasar de un objeto a una clase total de objetos que contiene al primero.

## 3. Especialización

Especialización es pasar de la consideración de una serie determinada de objetos a la de una serie más pequeña contenida en la primera. Por ejemplo, especializamos cuando pasamos de la consideración de polígonos a la de polígonos regulares, y especializamos aún más cuando pasamos de polígonos regulares con $n$ lados al triángulo equilátero.

Estos dos pasos subsecuentes fueron efectuados de dos maneras típicamente diferentes. En el primero, de polígonos a polígonos regulares introducimos una restricción, esto es, que todos los lados y los ángulos del polígono son iguales. En el segundo paso sustituimos un objeto variable por uno determinado: pusimos 3 donde estaba el entero variable $n$.

Con mucha frecuencia especializamos al pasar de una clase total de objetos a un objeto contenido en ella. Por ejemplo, cuando deseamos comprobar alguna aserción general sobre números primos tomamos algún número primo, digamos el 17, y observamos si esa aserción general es verdadera o no para este número.

## 4. Analogía

No hay nada vago o cuestionable en los conceptos de generalización y especialización. Sin embargo, cuando empezamos a discutir la analogía nos movemos sobre fundamento menos sólido.

*Analogía* es una especie de semejanza. Es, diríamos, semejanza sobre un nivel definido y conceptual. Aún podemos expresarnos con más exactitud. La diferencia esencial entre la analogía y otras clases de semejanza yace, en mi opinión, en las intenciones del pensador. Objetos semejantes son aquellos que concuerdan entre sí en algún aspecto. Si usted trata de delimitar el aspecto en que concuerdan para definir conceptos, usted mira estos objetos semejantes como *análogos*. Si tenemos éxito en aclarar los conceptos, hemos *clarificado* la analogía.

Comparando una mujer joven a una flor, los poetas sienten alguna semejanza, aunque normalmente ellos no contemplan la analogía. De hecho, apenas intentan abandonar el nivel emocional o reducir la comparación a algo mensurable o conceptualmente definible.

Mirando en un Museo de Historia Natural los esqueletos de varios mamíferos podemos encontrarlos a todos espantosos. Si ésta es toda la semejanza que podemos encontrar entre ellos, no veremos mucha analogía. Sin embargo, podemos encontrar una maravillosa y sugestiva analogía si nos fijamos en la mano de un hombre, las uñas de un gato, la pezuña de un caballo, las barbas de una ballena y las alas de un murciélago, órganos todos tan diferentemente usados como compuestos de partes semejantes y referidas unas a otras de modo semejante.

El último ejemplo ilustra el caso más típico de analogía clarificada; dos *sistemas* son análogos si *concuerdan en relaciones claramente definibles de sus partes respectivas.*

Por ejemplo, un triángulo en una superficie es análogo a un tetraedro en el espacio. En la superficie, dos líneas rectas no pueden constituir una figura finita, pero tres pueden constituir un triángulo. En el espacio, tres superficies no constituyen una figura finita, pero cuatro constituirán un tetraedro. La relación del triángulo a la superficie es la misma que la del tetraedro al espacio en cuanto que ambos están limitados por el número mínimo de elementos. He aquí la analogía:

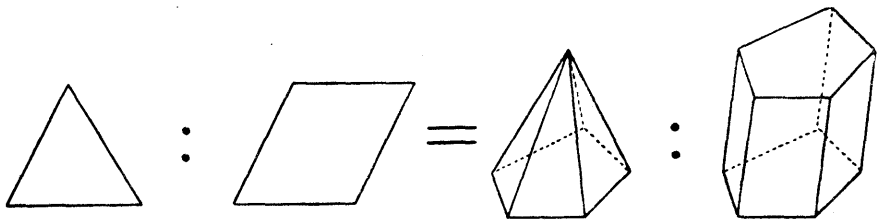


Fig. 2.1. Relaciones análogas en la superficie y el espacio.

Uno de los significados de la palabra griega «analogía», de donde proviene la española, es el de «proporción». De hecho, el sistema de los dos números 6 y 9 es «análogo» al sistema del 10 y el 15 en cuanto que ambos concuerdan en la razón de sus correspondientes términos,

$$6 : 9 = 10 : 15.$$

La proporcionalidad, o acuerdo en las razones de las partes correspondientes, que podemos ver intuitivamente en figuras geométricas semejantes, es un caso muy sugestivo de analogía.

He aquí otro ejemplo. Podemos considerar también al triángulo y la pirámide como figuras análogas. De un lado, tomemos un segmento de línea recta; por otro, un polígono. Conectemos todos los puntos del segmento con un punto exterior a la línea del segmento y obtendremos un triángulo. Conectemos todos los puntos del polígono con un punto exterior a la superficie del polígono y obtendremos una pirámide. De la misma manera podemos considerar un paralelogramo y un prisma como figuras análogas. En efecto, movamos un segmento o un polígono paralelo a sí mismo, a través de su línea o superficie, y uno describirá un paralelogramo, mientras el otro un prisma. Podemos también ser tentados a expresar las relaciones correspondientes entre superficies y cuerpos sólidos por una especie de proporción, y si, por una vez, no resistimos a la tentación, llegaremos a la figura 2.1. Esta figura modifica el significado usual de ciertos símbolos (: e =) de la misma manera que el significado de la palabra «analogía» fue modificado en el curso de la historia lingüística: de «proporción» a «analogía».

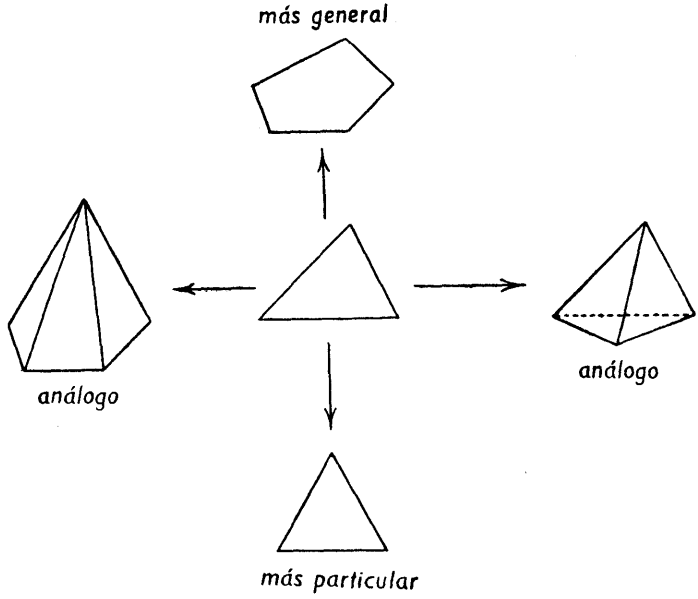



Fig. 2.2. Generalización, especialización, analogía.

El último ejemplo es aún instructivo en otro aspecto. La analogía, sobre todo si no ha sido aclarada, puede ser ambigua. Así, comparando las geometrías plana y sólida encontramos que un triángulo en el plano es análogo a un tetraedro en el espacio y también que un triángulo es

análogo a una pirámide. Ahora bien, ambas analogías son razonables y cada una valiosa en su aspecto. Hay varias analogías entre geometría sólida y plana, y no una sola privilegiada.

La figura 2.2 exhibe cómo, empezando con un triángulo, podemos ascender a un polígono por generalización, descender a un triángulo equilátero por especialización, o pasar a diferentes cuerpos sólidos por analogía; hay analogías a ambos lados.

Y acordémonos de no descuidar las analogías vagas. Pero si deseamos hacerlas estimables, intentemos clarificarlas.

## 5. Generalización, especialización y analogía

A menudo cooperan las tres para resolver problemas matemáticos *. Tomemos como ejemplo la prueba del teorema más conocido de la geometría elemental: el teorema de Pitágoras. La prueba que discutiremos no es nueva; la debemos al mismo Euclides (Euclides, VI, 31).

(1) Consideremos un triángulo rectángulo de lados $a$, $b$ y $c$, de los cuales el primero, $a$, es la hipotenusa. Nos proponemos mostrar que

$$a^2 = b^2 + c^2. \qquad \text{(A)}$$

Este punto sugiere que describamos los cuadros sobre los tres lados del triángulo rectángulo. Y así llegamos a la familiar parte I de nuestra figura compuesta: figura 2.3. (El lector deberá dibujar por su cuenta la figura para que aprenda su formación.)

(2) Los descubrimientos, aunque sean modestos, necesitan de observaciones: el reconocimiento de alguna relación. Nosotros podemos descubrir la prueba observando la *analogía* entre la familiar parte I de la figura y la no menos familiar parte II: el mismo triángulo que se levanta en I es dividido en II en dos partes por la altura perpendicular a la hipotenusa.

(3) Quizá no acertemos a percibir la analogía entre I y II. Esta analogía puede, sin embargo, ser hecha explícita por una *generalización* común de I y II, que es expresada en III. Encontramos allí otra vez el mismo triángulo rectángulo, y sobre sus tres lados se dibujan tres polígonos semejantes entre sí, aunque arbitrariamente diferentes.

(4) El área del cuadrado descrito sobre la hipotenusa en I es $a^2$.

---

* Esta sección reproduce una nota del autor en la *American Mathematical Monthly*, V, 55 (1948), págs. 241-243.

El área del cuadrado irregular descrito sobre la hipotenusa en III puede ser considerado igual $\lambda a^2$; el factor $\lambda$ es determinado como la razón de dos áreas dadas. Por ello, de la semejanza entre los tres polígonos descritos sobre los lados *a*, *b* y *c* en III se sigue que sus áreas serán iguales a $\lambda a^2$, $\lambda b^2$ y $\lambda c^2$, respectivamente.

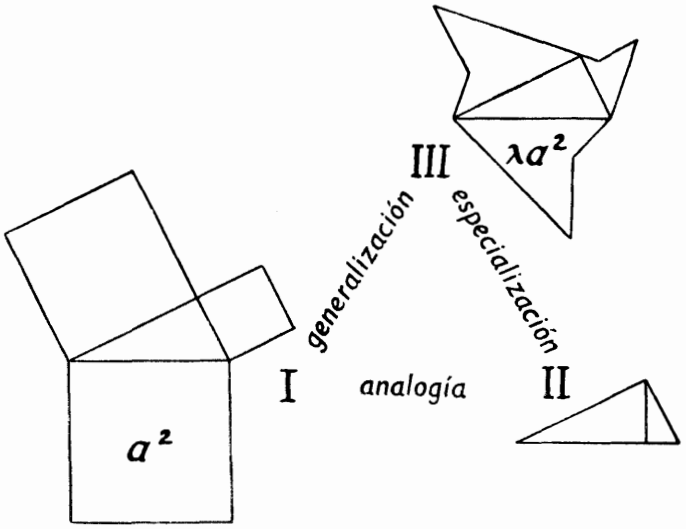



Fig. 2.3.

Ahora bien, si la ecuación (A) fuese cierta (como establece el teorema que deseamos probar), también sería cierta la siguiente ecuación:

$$\lambda a^2 = \lambda b^2 + \lambda c^2. \qquad \text{(B)}$$

De hecho, se necesita muy poca álgebra para derivar (B) de (A). Ahora bien, (B) representa una *generalización* del original teorema de Pitágoras: *Si tres polígonos semejantes son descritos sobre los tres lados de un triángulo rectángulo, el área descrito sobre la hipotenusa es igual a la de la suma de los otros dos.*

Resulta instructivo pensar que esta generalización es *equivalente* al caso con que empezamos. Efectivamente, nosotros podemos derivar las ecuaciones (A) y (B) una de otra, multiplicando o dividiendo por $\lambda$ (que, como razón de dos áreas, es diferente de 0).

(5) El teorema general expresado por (B) es equivalente no sólo al caso particular (A), sino a cualquier otro caso particular. Por tanto, si un tal caso particular resultase ser obvio, el caso general quedaría demostrado.

Ahora bien, al intentar *especializar* debemos buscar un caso particular conveniente. En realidad, II representa un caso tal. Pues el triángulo rectángulo descrito sobre su hipotenusa es semejante a los otros dos descritos a cada uno de sus lados, como es conocido y fácil de comprobar. Y, naturalmente, el área del triángulo total es igual a la suma de sus dos partes. Así, el teorema de Pitágoras resulta probado.

El razonamiento precedente es muy instructivo. Todo caso en el que aprendemos algo aplicable a otros casos resulta instructivo, y más instructivo cuanto más amplio el campo de aplicaciones posibles. Del ejemplo precedente hemos aprendido el uso de operaciones fundamentales, como la generalización, la especialización y la percepción de analogías. No hay, probablemente, ningún descubrimiento, ni en matemática avanzada o elemental, ni en cualquier otro tema, que pueda hacerse sin estas operaciones, sobre todo sin analogía.

El ejemplo anterior nos muestra cómo podemos ascender por generalización de un caso particular; cómo del representado en I pasamos a otra situación más general: la de III; y luego, cómo descendemos por especialización a un caso análogo: el de II. Ello nos muestra que el caso general puede ser lógicamente equivalente al caso particular, hecho usual en matemáticas, aunque sorprendente para el principiante o para el filósofo que se toma a sí mismo por avanzado. Nuestro ejemplo exhibe, finalmente, de modo sugestivo, cómo la generalización, la especialización y la analogía pueden combinarse en un esfuerzo por alcanzar la solución deseada. Observemos que sólo un mínimo de conocimiento preliminar es necesario para comprender del todo el razonamiento anterior.

## 6. Descubrimientos por analogía

La analogía parece tener participación en todos los descubrimientos, y en algunos se lleva la parte del león. Me gustaría ilustrar esto con un ejemplo que no fuese elemental, pero de interés histórico y bastante más impresionante que cualquier ejemplo elemental que yo pueda sacarme de la cabeza.

Jacques Bernouille, un matemático suizo (1654-1705), contemporáneo de Newton y Leibniz, descubrió la suma de varias series infinitas, pero fracasó en encontrar la suma de los cuadrados recíprocos,

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{25} + \frac{1}{36} + \frac{1}{49} + \cdots$$

«Si alguien encuentra —escribió Bernouille— lo que hasta ahora ha resistido nuestros esfuerzos y nos lo comunica le estaremos muy agradecidos por ello.»

El problema llamó la atención de otro matemático suizo, Leonhard Euler (1707-1783), que nació en Basilea, como Jacques Bernouille, y fue alumno del hermano de éste, Jean Bernouille (1667-1748). El encontró varias series para el número deseado (integrales definidas, otras series), ninguna de las cuales le satisfizo; usó estas expresiones para calcular la suma de siete cifras (1.644934). Sin embargo, esto era un valor aproximado y su meta era un valor exacto. Llegó a descubrirlo por casualidad. La analogía le había conducido a una conjetura extremadamente osada.

(1) Empecemos por repasar algunos elementos de álgebra esenciales al descubrimiento de Euler. Si la ecuación de $n$ grados

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n = 0$$

tiene $n$ diferentes raíces

$$\alpha_1, \alpha_2, \ldots, \alpha_n$$

el polinomio de la izquierda puede ser representado como el producto de $n$ factores lineales,

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n =$$
$$a_n (x - \alpha_1) (x - \alpha_2) \ldots (x - \alpha_n).$$

Comparemos los términos con el mismo valor de $x$ a ambos lados de esta identidad y derivemos las conocidas relaciones entre las raíces y los coeficientes de una ecuación, la más sencilla de las cuales es

$$a_{n-1} = -a_n(\alpha_1 + \alpha_2 + \ldots + \alpha_n);$$

lo que encontramos comparando los términos con $x^{n-1}$.

Hay otra manera de representar la descomposición en factores lineales. Si ninguna de las raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$ es igual a 0, o (lo que es lo mismo) si $a_0$ es diferente de 0, tenemos también

$$a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

$$= a_0 \left(1 - \frac{x}{\alpha_1}\right)\left(1 - \frac{x}{\alpha_2}\right) \ldots \left(1 - \frac{x}{\alpha_n}\right)$$

y

$$a_1 = -a_0 \left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \ldots + \frac{1}{\alpha_n}\right).$$

Hay aún otra variante. Supongamos que la ecuación es de $2n$ grados y tiene la forma

$$b_0 - b_1x^2 + b_2x^4 - \ldots + (-1)^n b_n x^{2n} = 0$$

y $2n$ diferentes raíces

$$\beta_1, -\beta_1, \beta_2, -\beta_2, \quad \ldots, \quad \beta_n, \quad -\beta_n.$$

Entonces,

$$b_0 - b_1x^2 + b_2x^4 - \ldots + (-1)^n b_n x^{2n}$$

$$= b_0 \left(1 - \frac{x^2}{\beta_1^2}\right)\left(1 - \frac{x^2}{\beta_2^2}\right) \ldots \left(1 - \frac{x^2}{\beta_n^2}\right)$$

y

$$b_1 = b_0 \left(\frac{1}{\beta_1^2} + \frac{1}{\beta_2^2} + \ldots + \frac{1}{\beta_n^2}\right).$$

(2) Euler considera la ecuación

$$\text{sen } x = 0$$

o

$$\frac{x}{1} - \frac{x^3}{1\cdot 2\cdot 3} + \frac{x^5}{1\cdot 2\cdot 3\cdot 4\cdot 5} - \frac{x^7}{1\cdot 2\cdot 3\cdots 7} + \ldots = 0.$$

El lado de la izquierda tiene infinidad de términos, es de «grado infinito». Por tanto, no es extraño, dice Euler, que haya una infinidad de raíces

$$0, \quad \pi, \quad -\pi, \quad 2\pi, \quad -2\pi, \quad 3\pi, \quad -3\pi, \quad \ldots$$

Euler descarta la raíz 0. Luego divide el lado izquierdo de la ecuación por $x$, el factor lineal correspondiente a la raíz 0, y obtiene así la ecuación

$$1 - \frac{x^2}{2\cdot 3} + \frac{x^4}{2\cdot 3\cdot 4\cdot 5} - \frac{x^6}{2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7} + \ldots = 0$$

con las raíces

$$\pi, -\pi, \quad 2\pi, -2\pi, 3\pi, -3\pi, \quad \ldots$$

Hemos visto una situación análoga, bajo (1), cuando discutimos la última variante de descomposición en factores lineales. Euler concluye, por analogía, que

$$\frac{\text{sen } x}{x} = 1 - \frac{x^2}{2 \cdot 3} + \frac{x^4}{2 \cdot 3 \cdot 4 \cdot 5} - \frac{x^6}{2 \cdot 3 \cdot \cdots 7} + \cdots$$

$$= \left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right) \cdots,$$

$$\frac{1}{2 \cdot 3} = \frac{1}{\pi^2} + \frac{1}{4\pi^2} + \frac{1}{9\pi^2} + \cdots,$$

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots \quad \frac{\pi^2}{6}.$$

Es decir, la serie que resistió los esfuerzos de Jacques Bernouille; una osada conclusión.

(3) Euler sabía muy bien que su conclusión era osada. «El método era nuevo y nunca usado para un propósito semejante», escribió diez años más tarde. El mismo se dio cuenta de algunas objeciones y otras muchas le fueron propuestas por sus amigos matemáticos, una vez recobrados de la primera sorpresa.

Sin embargo, Euler tenía razones para confiar en su descubrimiento. En primer lugar, el valor numérico de la suma de las series que él había calculado antes estaba de acuerdo hasta el fin con $\pi^2/6$. Al comparar los demás coeficientes de su expresión sen $x$ como producto encontró la suma de otras series dignas de ser tenidas en cuenta, como la de los recíprocos de cuatro valores,

$$1 + \frac{1}{16} + \frac{1}{81} + \frac{1}{256} + \frac{1}{625} + \cdots = \frac{\pi^4}{90}.$$

De nuevo examinó los valores numéricos y de nuevo los encontró de acuerdo.

(4) Euler también comprobó su método con otros ejemplos. Al hacerlo así acertó volviendo a derivar la suma $\pi^2/6$ de la serie de Jacques Bernouille con varias modificaciones a su primer aproximación. También acertó a redescubrir con su método la suma de importantes series debidas a Leibniz.

Discutamos el último punto, considerando, según hace Euler, la ecuación

$$1 - \operatorname{sen} x = 0,$$

que tiene como raíces

$$\frac{\pi}{2}, \quad -\frac{3\pi}{2}, \quad \frac{5\pi}{2}, \quad -\frac{7\pi}{2}, \quad \frac{9\pi}{2}, \quad -\frac{11\pi}{2}, \quad \ldots$$

Cada una de estas raíces es, sin embargo, una raíz doble. (La curva $y = \operatorname{sen} x$ no intercede la línea $y = 1$ en estas abscisas, sino que es tangente a ella. La derivada del lado izquierdo se desvanece para los mismos valores de $x$, pero no así la segunda derivada). Por tanto, la ecuación

$$1 - \frac{x}{1} + \frac{x^3}{1 \cdot 2 \cdot 3} - \frac{x^5}{1 \cdot 2 \cdot 3 \cdot 5} + \ldots = 0$$

tiene las ra´ces

$$\frac{\pi}{2}, \quad \frac{\pi}{2}, \quad -\frac{3\pi}{2}, \quad -\frac{3\pi}{2}, \quad \frac{5\pi}{2}, \quad \frac{5\pi}{2}, \quad -\frac{7\pi}{2}, \quad -\frac{7\pi}{2}, \quad \ldots$$

y la conclusión analógica de Euler conduce a la descomposición en factores lineales

$$1 - \operatorname{sen} x = 1 - \frac{x}{1} + \frac{x^3}{1 \cdot 2 \cdot 3} - \frac{x^5}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} + \ldots$$

$$= \left(1 - \frac{2x}{\pi}\right)^2 \left(1 + \frac{2x}{3\pi}\right)^2 \left(1 - \frac{2x}{5\pi}\right)^2 \left(1 + \frac{2x}{7\pi}\right)^2 \ldots$$

Comparando el coeficiente de $x$ en ambos lados, obtenemos

$$-1 = -\frac{4}{\pi} + \frac{4}{3\pi} - \frac{4}{5\pi} + \frac{4}{7\pi} - \ldots,$$

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \ldots$$

Esta es la célebre serie de Leibniz; el osado procedimiento de Euler conduce a un resultado conocido. «Nuestro método, que puede parecerle

a alguien no muy convincente, alcanza así una gran confirmación. En consecuencia, por lo que a nosotros se refiere, nos sentimos incapaces de dudar de las demás cosas que han sido derivadas por el mismo método.»

(5) Sin embargo, Euler permaneció en duda. Y continuó las verificaciones numéricas descritas arriba (3), examinó más series y más grados decimales, hallando acuerdo en todos ellos. Intentó otras aproximaciones también y, finalmente, logró verificar, no sólo numéricamente, sino exactamente, el valor $\pi^2/6$ para la serie de Bernouille. Encontró una prueba que, aunque oculta e ingeniosa, estaba basada en consideraciones más normales y fue aceptada como completamente rigurosa. Así, la más conspicua consecuencia del descubrimiento de Euler fue satisfactoriamente verificada.

Estos argumentos parece que convencieron a Euler de que su resultado era correcto *.

## 7. Analogía e inducción

Nosotros deseamos aprender algo de la naturaleza del razonamiento inventivo e inductivo. ¿Qué podemos aprender de la historia precedente?

(1) El paso decisivo de Euler fue osado. En lógica estricta fue una falacia indebida: aplicó una regla a un caso para el que la regla no estaba hecha, una regla sobre ecuaciones algebraicas a una ecuación que no era algebraica. Lógicamente, el paso de Euler no estaba justificado. Sin embargo, la analogía lo justificaba; analogía que tiene en su haber los mejores logros de una ciencia en crecimiento a la que algunos años después él mismo llamó «Análisis del infinito». Otros matemáticos, anteriores a Euler, habían pasado de las diferencias finitas a las infinitamente pequeñas, de las sumas con número finito de cifras a aquellas con número infinito, de los productos finitos a los infinitos. Y así Euler pasó de las ecuaciones de grado finito (ecuaciones algebraicas) a las de grado infinito, aplicando a lo infinito las reglas hechas para lo finito.

Esta analogía, este paso de lo finito a lo infinito, está cubierto de añagazas. ¿Cómo las evitó Euler? «Era un genio», han contestado algunas personas, pero ello no supone aplicación ninguna. Euler tenía

---

* Mucho más tarde, casi diez años después de su primer descubrimiento, Euler volvió sobre el tema, contestó objeciones, completó con alguna extensión su original aproximación heurística y dio una nueva y esencialmente diferente prueba. Véase L. Euler, *Opera Omnia*, ser. 1, vol. 14, págs. 73-86, 138-155, 177-186 y también páginas 156-176, que contiene una nota de Paul Stäckel sobre la historia del problema.

serias razones para confiar en su descubrimiento. Nosotros podemos comprender sus razones con un poco de sentido común, sin ninguna milagrosa perspicacia específica del genio.

(2) Las razones de Euler para confiar en su descubrimiento, resumidas en lo anterior *, *no* son demostrativas. Euler no reexamina los fundamentos de su conjetura ** para el osado paso de lo finito a lo infinito; él examina sólo las consecuencias y mira la verificación de tales consecuencias como argumentos en pro de su conjetura; acepta las verificaciones aproximativas y las exactas, aunque parece dar más peso a las segundas. Examina también las consecuencias de conjeturas *** análogas íntimamente relacionadas y mira la verificación de tales consecuencias con argumentos para su conjetura.

Las razones de Euler son, de hecho, inductivas. Utiliza un procedimiento típicamente inductivo para examinar las consecuencias de una conjetura y juzgarlas sobre la base de tal examen. En la investigación científica, como en la vida ordinaria, nosotros creemos, o deberíamos creer, las conjeturas más o menos según el mayor o menor acuerdo con los hechos de sus consecuencias observables.

En definitiva, Euler parece pensar de la misma manera que las gentes razonables piensan, sean o no científicas. El parece aceptar algunos principios: *Una conjetura adquiere más crédito con la verificación de una nueva consecuencia.* Y: *Una conjetura alcanza más crédito si una conjetura análoga adquiere mayor crédito.*

¿Son de este tipo los principios subyacentes al proceso de inducción?

## EJEMPLOS Y COMENTARIOS

### Primera parte

**1.** *La generalización correcta.* A. Encontramos tres números, $x$, $y$ $z$, que satisfacen el siguiente sistema de ecuaciones:

$$\begin{aligned} 9x - 6y - 10z &= 1, \\ -6x + 4y + 7z &= 0, \\ x^2 + y^2 + z^2 &= 9. \end{aligned}$$

---

* Bajo la sección 6 (3), (4), (5). Para el resumen del propio Euler véase *Opera Omnia*, ser. 1, vol. 14, pág. 140.

** La representación de sen $x$ es un producto infinito.

*** Especialmente el producto para 1 — sen $x$.

Si tiene usted que resolver A, ¿cuál de las tres generalizaciones siguientes le parece la más valiosa sugerencia, B, C o D?

B. Hallar tres incógnitas de un sistema de tres ecuaciones.

C. Hallar tres incógnitas de un sistema de tres ecuaciones, de las cuales las dos primeras son lineales y la tercera cuadrática.

D. Hallar $n$ incógnitas de un sistema de $n$ ecuaciones, de las cuales la primera $n - 1$ es lineal.

**2.** Un punto y una pirámide «regular» con base hexagonal son dados en posición. (Se denomina «regular» a una pirámide cuya base es un polígono regular, el centro del cual es el pie de la altura de la pirámide.) Hallar un plano que pase por el punto dado y biseccione el volumen de la pirámide dada.

Con objeto de ayudarle le haré la siguiente pregunta: ¿Cuál es la generalización correcta?

**3.** A. Tres líneas rectas que no están en el mismo plano pasan por el mismo punto $O$. Hallar un plano que pasando por $O$ esté inclinado igualmente a las tres líneas.

B. Tres líneas rectas que no están en el mismo plano pasan por el mismo punto. El punto $P$ está en una de las líneas; hallar el plano, que, pasando por $P$, esté igualmente inclinado a las tres líneas.

Compare los problemas A y B. ¿Podría usted utilizar la solución de uno para resolver el otro? ¿Cuál es su conexión lógica?

**4.** A. Calcule la integral

$$\int_{-\infty}^{\infty} (1 + x^2)^{-3}\, dx.$$

B. Calcule la integral

$$\int_{-\infty}^{\infty} (p + x^2)^{-3}\, dx,$$

en la cual $p$ es un número positivo determinado.

Compare los problemas A y B. ¿Podría usted utilizar la solución de uno para resolver el otro? ¿Cuál es su conexión lógica?

**5.** *Un caso particular extremo.* Hay dos hombres sentados en una mesa de forma rectangular. Uno coloca una moneda sobre la mesa, después el otro hace lo mismo, y así alternativamente. Se entiende que cada moneda yace sobre la mesa sin que ninguna esté montada sobre

otra. El jugador que coloca la última moneda se lleva todo el dinero. ¿Quién ganará, teniendo en cuenta que cada uno de ellos ha hecho el mejor juego posible?

He aquí un acreditado rompecabezas. Yo tuve la oportunidad de observar a un distinguido matemático cuando se le propuso el problema. El empezó diciendo: «Supongamos que la mesa es tan pequeña que resulta cubierta por un penique; entonces es obvio que el primer jugador debe ganar». Es decir, empezó por entresacar un *caso particular extremo* en que la solución resulta obvia.

Desde este caso particular podemos llegar a una solución completa si imaginamos la mesa extenderse gradualmente para alojar más y más monedas. Aún parece ser mejor *generalizar* el problema y pensar en mesas de varias formas y tamaños. Si usted observa que la mesa tiene un centro de simetría y que la *generalización correcta* debe considerar mesas con un centro de simetría, entonces usted ha alcanzado la solución o está muy cerca de alcanzarla.

**6.** *Construir una tangente común a dos círculos dados.* Con objeto de ayudarle le haré la siguiente pregunta: ¿Hay un caso particular extremo más accesible?

**7.** *Un caso particular guía.* El área de un polígono es $A$ y su plano forma con un segundo plano el ángulo $\alpha$. El polígono es proyectado ortogonalmente sobre el segundo plano. Hallar el área de la proyección.

Observemos que no nos es dada la forma del polígono. Por tanto, tenemos un sinfín de variedades posibles. ¿Qué forma adoptaremos? ¿Con cuál de ellas empezaremos?

Hay una forma particular de muy fácil manejo: el rectángulo, cuya base es paralela a la línea $l$, intersección del plano de la figura proyectada con el de la proyección. Si la base de tal rectángulo es $a$ y su altura $b$, el área será $ab$ y las cantidades correspondientes a la proyección, $a$, $b$, $\cos\alpha$ y $ab\cos\alpha$. El área del rectángulo sabemos que es $A$, la de su proyección será $A\cos\alpha$.

Este caso particular del rectángulo con base paralela a $l$ es no sólo especialmente accesible, sino un *caso particular guía*. Los otros casos siguen, pero *la solución del problema en el caso particular guía envuelve la solución del caso general*. De hecho, al empezar con el rectángulo de base paralela a $l$ podemos extender la regla «área de proyección $A\cos\alpha$» sucesivamente a todas las otras figuras. Primero, a los triángulos rectángulos con un pie paralelo a $l$ (por bisección del rectángulo con que empezamos); después, a todos los triángulos con un lado paralelo a $l$ (por combinación de dos triángulos rectángulos); finalmente, a un polí-

gono general (analizando los triángulos mencionados). Igualmente podríamos pasar a figuras con límites curvilíneos (considerándolos como límites de polígonos).

**8.** El ángulo en el centro de un círculo es doble al ángulo en la circunferencia de la misma base, esto es, del mismo arco. (EUCLIDES, III, 20.)

Aunque el ángulo del centro nos sea dado, el ángulo de la circunferencia no está determinado todavía, pues admite varias posiciones. En la prueba usual del teorema (prueba de Euclides) ¿cuál es la «posición particular guía»?

**9.** El teorema de Cauchy, fundamental en la teoría de las funciones analíticas, asegura que la integral de una función tal se desvanece a lo largo de una arbitraria curva cerrada en cuyo interior la función es regular. Nosotros podemos considerar el caso particular del teorema de Cauchy, en el que la curva cerrada es un triángulo como un caso particular guía: una vez probado el teorema para un triángulo podemos con facilidad extenderlo sucesivamente a polígonos (combinando triángulos) y a curvas (considerándolas como límites de polígonos). Observemos la analogía con los ejemplos 7 y 8.

**10.** *Un caso particular representativo.* Tenemos que resolver un problema de polígonos con $n$ lados. Dibujemos un pentágono y resolvamos el problema para él, estudiamos después la solución y observamos que vale igualmente para el caso general, de polígonos de $n$ lados, en el caso particular de $n = 5$. Entonces podemos llamar a $n = 5$ un caso particular *representativo*, puesto que representa el caso general. Por supuesto, el objeto de ser realmente representativo el caso $n = 5$ no debe tener ninguna interpretación especial que nos extravíe. El caso particular representativo *no* deberá ser más sencillo que el caso general.

Casos particulares representativos son a menudo convenientes en la enseñanza. Así podemos resolver un problema de determinantes con $n$ filas discutiendo cuidadosamente un determinante con sólo tres.

**11.** *Un caso análogo.* Se presenta el problema de diseñar aviones en los que el peligro de fractura de cráneo en caso de accidente se reduzca al mínimo. Un médico, para estudiar el problema, experimenta con huevos que rompe bajo distintas condiciones. ¿Qué es lo que hace? Ha *modificado* el problema original, y estudiando ahora un *problema auxiliar* rompe huevos en lugar de cráneos. El lazo entre ambos problemas, el original y el auxiliar, es la *analogía*. Desde un punto de vista mecánico, una cabeza de hombre y un huevo de gallina son grosera-

mente análogos: ambos consisten en una concha rígida y frágil que contiene materia gelatinosa.

**12.** Si dos líneas rectas en una superficie son cortadas por tres planos paralelos, los segmentos resultantes son proporcionales.

Al objeto de ayudarle a encontrar una prueba le presento la siguiente pregunta: ¿Hay un teorema análogo más sencillo?

**13.** Las cuatro diagonales de un paralelepípedo tienen un punto común que es el punto medio de todas ellas.

¿Hay un teorema análogo más sencillo?

**14.** La suma de dos ángulos frontales de un triedro es mayor que el tercer ángulo frontal.

¿Hay un teorema análogo más sencillo?

**15.** Consideremos el tetraedro como el cuerpo sólido más análogo al triángulo. Enumerar los conceptos de la geometría sólida que son análogos a los siguientes conceptos de la geometría plana: *paralelogramo, rectángulo, cuadrado, bisectriz de un ángulo.* Establecer un teorema de la geometría sólida que sea análogo al siguiente teorema de la geometría plana: *Las bisectrices de los tres ángulos de un triángulo se unen en un punto que es el centro del círculo inscrito en el triángulo.*

**16.** Consideremos una pirámide como cuerpo sólido análogo a un triángulo. Enumerar los sólidos que son análogos a las siguientes figuras planas: *paralelogramo, rectángulo, círculo.* Establecer un teorema de la geometría sólida que es análogo al siguiente teorema de la geometría plana: *El área de un círculo es igual al área de un triángulo, cuya base tiene la misma longitud que el perímetro del círculo y cuya altura es igual al radio.*

**17.** Invéntese un teorema de la geometría sólida que sea análogo al siguiente teorema de la geometría plana: *La altura de un triángulo isósceles pasa por el punto medio de la base.*

¿Qué figura sólida debemos considerar como análoga a un triángulo isósceles?

**18.** *Grandes analogías.* (1) Los anteriores ejemplos 12-17 insisten sobre la analogía entre *geometría sólida* y *geometría plana.* Esta analogía tiene muchos aspectos y es, por ello, a menudo, ambigua y no siempre aclaradora; pero es una fuente inagotable de nuevas sugerencias y descubrimientos.

(2) Los números y las figuras no son sólo objeto de la matemática, en cuanto que la matemática es básicamente inseparable de la lógica

y se relaciona con todos los objetos que son capaces de una teoría exacta. Números y figuras son, sin embargo, el más normal objeto de la matemática, y al matemático le gusta ilustrar los números con propiedades de las figuras y las figuras con propiedades de los números. De aquí surgen innumerables aspectos de la analogía entre *números* y *figuras*. Algunos de estos aspectos están muy claros. Así, en geometría analítica estudiamos correspondencias bien definidas entre objetos y relaciones algebraicas y geométricas. Sin embargo, la variedad de las figuras geométricas es inagotable y lo mismo la variedad de operaciones posibles con los números; de aquí que sean posibles las correspondencias entre estas variedades.

(3) El estudio de los límites y sus procesos introduce otra clase de analogía, que podemos llamar la analogía entre lo *finito* y lo *infinito*. Así, las series infinitas e integrales son análogas de diversos modos a las sumas finitas, cuyos límites son: el cálculo diferencial es análogo al cálculo de diferencias finitas; las ecuaciones diferenciales, especialmente las ecuaciones lineales y las diferenciales homogéneas, son de alguna manera análogas a las ecuaciones algebraicas, y así sucesivamente. Una importante rama de la matemática, y relativamente reciente, es la teoría de las ecuaciones integrales, que da una sorprendente y bella respuesta a la cuestión: ¿Cuál es, en el cálculo integral, el análogo de un sistema de $n$ ecuaciones lineales con $n$ incógnitas? La analogía entre infinito y finito es muy interesante por sus particulares dificultades y trampas. Puede llevar al descubrimiento o al error; ver ejemplo 46.

(4) Galileo descubrió la trayectoria parabólica de los proyectiles y las leyes cuantitativas de su movimiento; fue también un gran descubridor en la astronomía. Con su recién inventado telescopio descubrió los satélites de Júpiter. Observó que estos satélites giran en torno a Júpiter de modo análogo a como la Luna gira sobre la Tierra y los planetas alrededor del Sol. Aún más: descubrió las fases de Venus y observó su semejanza con las fases de la Luna. Estos descubrimientos son recibidos como una gran confirmación de la teoría heliocéntrica de Copérnico, calurosamente discutida en aquel tiempo. Es curioso que Galileo no percibiese la analogía entre el movimiento de los cuerpos celestes y el movimiento de los proyectiles, que puede verse intuitivamente. La trayectoria de un proyectil muestra su lado cóncavo hacia la Tierra, igual que la de la Luna. Newton insistió sobre esta analogía: «... una piedra arrojada al espacio es obligada, por la presión de su propio peso, a abandonar la trayectoria rectilínea que había seguido por el impulso inicial, describiendo una línea curva en el aire, y ... al fin,

cae en el suelo; y cuanto mayor es la velocidad con que es proyectada más lejano es el lugar en que cae a la Tierra. De manera que si suponemos que la velocidad se va incrementando, hasta describir arcos de 1, 2, 5, 10, 100, 1.000 millas antes de caer a tierra, llegará un momento en que excederá los límites de la misma Tierra y pasará al espacio sin tocarla» * Véase figura 2.4.

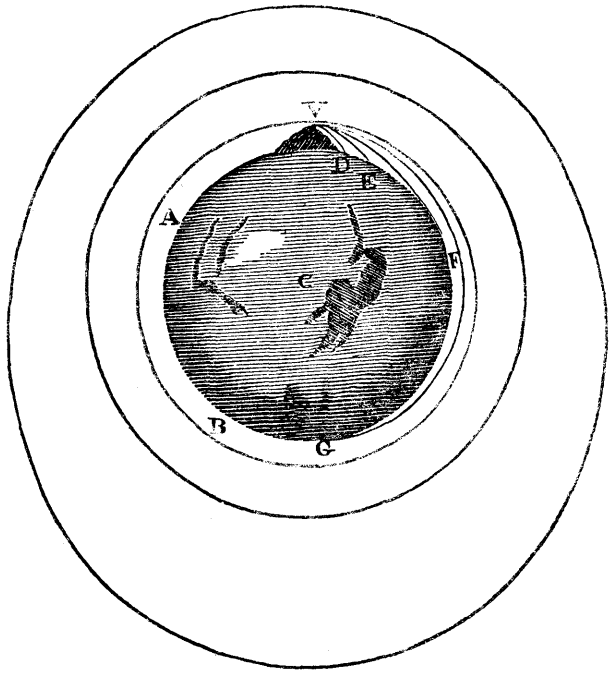



Fig. 2.4. De la trayectoria de la piedra a la de la Luna. Tomado de los *Principia*, de Newton.

Al variar continuamente la trayectoria de la piedra nos conduce a la de la Luna. Y así, piedra y Luna son a la Tierra como los satélites a Júpiter o Venus y los otros planetas al Sol. Sin visualizar esta analogía, nosotros podemos comprender, sólo de modo muy imperfecto, el descubrimiento de la gravitación universal, que podemos considerar todavía como uno de los más grandes descubrimientos científicos de todos los tiempos.

**19.** *Analogías clarificadas.* La analogía es, a menudo, vaga. La contestación a la pregunta qué es análogo a qué resulta ambigua con

* Sir Isaac Newton, *Mathematical Principles of Natural Philosophy and his System of the World*, traducido por Motte, revisado por Cajori, Berkeley, 1946; ver pág. 551.

frecuencia. La vaguedad de la analogía no debe disminuir su interés y su utilidad; sin embargo, aquellos casos en que el concepto de analogía atenta contra la claridad de los conceptos lógicos o matemáticos merece aún especial consideración.

(1) La analogía es semejanza de relaciones. Esta semejanza tiene un claro significado si *las relaciones están dirigidas por las mismas leyes.* En este sentido, la suma de números es análoga a la multiplicación, en tanto que suma y multiplicación están sujetas a las mismas reglas. Ambas son conmutativas y asociativas:

$$a + b = b + a, \qquad ab = ba,$$

$$(a + b) + c = a + (b + c), \qquad (ab)\, c = a\,(bc).$$

Ambas admiten también operaciones inversas; las ecuaciones

$$a + x = b, \quad ax = b$$

son semejantes en tanto que ambas admiten una solución y nada más que una. (Con objeto de establecer la regla sin excepciones debemos admitir números negativos cuando consideramos la suma, y excluir el caso $a = 0$ en la multiplicación.) En este contexto, resta es análoga a división; en efecto, las soluciones a las ecuaciones de arriba son

$$x = b - a, \qquad x = \frac{b}{a},$$

respectivamente. Por otro lado, el número 0 es análogo al 1; así, la suma de 0 a cualquier número, como la multiplicación por 1, no cambian el número

$$a + 0 = a, \quad a \cdot 1 = a.$$

Estas leyes son las mismas para varias clases de números; podemos considerar aquí números racionales, reales o complejos. En general, *sistemas de objetos sujetos a las mismas leyes* (o axiomas) *fundamentales* pueden ser considerados como análogos unos a otros, y esta clase de analogía tiene un significado muy claro.

(2) La suma de números *reales* es análoga a la de números *positivos* en un nuevo sentido. Cualquier número real $r$ es el logaritmo de algún número positivo $p$,

$$r = \log p.$$

(Si nos atenemos a logaritmos ordinarios, $r = -2$ si $p = 0.01$.) Por virtud de esta relación, a cada número positivo corresponde un número real perfectamente determinado y a cada número real otro positivo perfectamente determinado. En esta correspondencia la suma de números reales corresponde a la multiplicación de números positivos. Si

$$r = \log p, \quad r' = \log p', \quad r'' = \log p'',$$

entonces cualquiera de las relaciones siguientes implica la otra:

$$r + r' = r'', \quad pp' = p''.$$

La fórmula de la izquierda y la de la derecha nos dicen lo mismo, aunque en lenguaje diferente. Llamemos a uno de los números coordinados la traducción del otro; por ejemplo, llamemos al número real $r$ (el logaritmo de $p$) la *traducción* de $p$ y a $p$ el *original* de $r$. (Hemos intercambiado las palabras «traducción» y «original», pues debíamos elegir, pero una vez hecha la elección nos adherimos a ella.) En esta terminología la suma aparece como traducción de la multiplicación; la resta como traducción de la división; 0, como traducción de 1, las leyes conmutativas y asociativas para la suma de números reales son concebidas como la traducción de estas leyes para la multiplicación de números positivos. La traducción es, por supuesto, diferente del original, pero es una traducción correcta en el siguiente sentido: de una relación entre elementos originales podemos concluir con certeza la correspondiente relación entre los elementos traducidos, y *viceversa*. Tal traducción correcta, que es *una correspondencia de uno a uno preservando las leyes de ciertas relaciones*, la llamamos *isomorfismo* en el lenguaje técnico de los matemáticos. Isomorfismo es una especie de analogía completamente clarificada.

(3) Una tercera clase de analogía totalmente clarificada es la que los matemáticos llaman en lenguaje técnico *homomorfismo*. Nos llevaría demasiado tiempo estudiar un ejemplo con suficiente detalle, o dar una descripción exacta, pero podemos intentar comprender la siguiente descripción aproximada. Homomorfismo es una especie de *traducción sistemáticamente abreviada*. El original no es sólo traducido a otro lenguaje, sino también abreviado, de modo que el resultado final de traducción y abreviación es sistemáticamente condensado en un medio, tercio o cualquiera otra fracción de la extensión original. Las sutilezas pueden perderse en una abreviación semejante, pero todo lo que está en el original debe ser representado por algo en la traducción y, a escala reducida, las relaciones serán mantenidas.

**20.** *Algunas citas.* «Veamos si, por casualidad, podemos concebir algún otro problema general que contenga el original y sea más fácil de resolver. Así, cuando buscamos la tangente de un punto dado imaginamos buscar una línea recta que intercede con la curva dada en el punto dado. Una vez resuelto este problema, que es más fácil hacerlo por álgebra, encontramos el caso de la tangente como un caso particular, es decir, el caso particular en que la distancia dada es mínima y, reducida a un punto, se desvanece» (LEIBNIZ).

«Sucede, a menudo, que el problema general es más fácil que el particular, cuando atacamos éste directamente» (P. G. LEJEUNE-DIRICHLET, R. DEDEKIND).

«[Puede ser útil] reducir el género a varias especies, incluso a pocas especies. Sin embargo, lo más útil es reducir el género al mínimo de especies» (LEIBNIZ).

«Es propio de la filosofía considerar lo semejante, aun en cosas muy distantes unas de otras» (ARISTÓTELES).

«Las comparaciones son de gran valor en cuanto reducen las relaciones desconocidas a otras conocidas.

»La comprensión propia es, en definitiva, una trabazón de relaciones. Pero comprendemos más pura y distintamente una relación cuando la reconocemos como la misma en un número muy amplio de casos y entre objetos completamente heterogéneos» (ARTURO SCHOPENHAUER).

Sin embargo, hay dos clases de generalizaciones. Una de pacotilla, otra valiosa. Cuando generalizamos es fácil *diluir;* lo importante es *condensar.* Diluir un poco de vino en agua es fácil y barato. Preparar un extracto refinado y condensado con varios y buenos ingredientes es mucho más difícil, aunque más valioso. Generalizar por condensación supone reunir en un concepto con amplio horizonte de ideas diversas lo que antes estaba profusamente desparramado. Así, la teoría de los grupos reduce a una expresión común lo que antes estaba disperso en álgebra, teoría de los números, análisis, geometría, cristalografía y otros dominios. La otra generalización está más de moda hoy en día de lo que estuvo antiguamente. Suele diluir una pequeña idea en una gran terminología. El autor prefiere tomar esta pequeña idea de cualquiera, contenerse de añadir ninguna observación original y evitar resolver cualquier problema, excepto aquellos que levantan dificultades de su propia terminología. Sería muy fácil citar ejemplos, pero no quiero discutir con nadie *.

---

* Cf. G. POLYA y S. SZEGÖ, *Aufgaben und Lehrstäze aus der Analysis,* vol. 1, página VII.

## Segunda parte

Los ejemplos y comentarios de esta segunda parte están todos conectados con la sección 6 y ellos mismos entre sí. Muchos de ellos se refieren directa o indirectamente al ejemplo 21, que será el primero.

**21.** *La conjetura E.* Consideremos la ecuación

$$\operatorname{sen} x = x\left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right) \cdots$$

como una conjetura; lo llamaremos «la conjetura *E*». Siguiendo a Euler, deseamos investigar esta conjetura inductivamente.

La investigación inductiva de una conjetura requiere confrontar sus consecuencias con los hechos. A menudo «predecimos desde *E* y verificamos». «Predecir desde *E*» significa derivar bajo el supuesto de que *E* es cierto; «verificar» significa derivar sin esta presunción. Un hecho «concuerda con *E*» si podemos derivar de esa presunción que *E* es cierto.

En lo siguiente nosotros utilizaremos los elementos del cálculo (que, desde el punto de vista formal, eran conocidos por Euler en la época de su descubrimiento), incluyendo el riguroso concepto de límite (sobre el que Euler no llegó nunca a completa claridad). Por nuestra parte, utilizaremos los procedimientos de límites que estén justificados, aunque sin entrar en el detalle de su justificación.

**22.** Sabemos que $\operatorname{sen}(-x) = -\operatorname{sen} x$. ¿Está este hecho de acuerdo con *E*?

**23.** Predecir desde *E* y verificar el valor del producto infinito

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\left(1 - \frac{1}{16}\right) \cdots \left(1 - \frac{1}{n^2}\right) \cdots$$

**24.** Predecir desde *E* y verificar el valor del producto infinito

$$\left(1 - \frac{4}{9}\right)\left(1 - \frac{4}{16}\right)\left(1 - \frac{4}{25}\right) \cdots \left(1 - \frac{4}{n^2}\right) \cdots$$

**25.** Comparar los ejemplos 23 y 24, y generalizar.

**26.** Predecir desde $E$ el valor del producto infinito

$$\frac{2 \cdot 4}{3 \cdot 3} \cdot \frac{4 \cdot 6}{5 \cdot 5} \cdot \frac{6 \cdot 8}{7 \cdot 7} \cdot \frac{8 \cdot 10}{9 \cdot 9} \cdots$$

**27.** Mostrar que la conjetura $E$ es equivalente a la proposición

$$\frac{\operatorname{sen} \pi z}{\pi} = \lim_{n=\infty} \frac{(z+n) \ldots (z+1)\, z\, (z-1) \ldots (z-n)}{(-1)^n (n!)^2}$$

**28.** Sabemos que $\operatorname{sen}(x+\pi) = -\operatorname{sen} x$. ¿Este hecho concuerda con $E$?

**29.** El método de la sección 6 (2) conduce a esta conjetura:

$$\cos x = \left(1 - \frac{4x^2}{\pi^2}\right)\left(1 - \frac{4x^2}{9\pi^2}\right)\left(1 - \frac{4x^2}{25\pi^2}\right) \cdots$$

Mostrar que esto no es sólo análogo a la conjetura $E$, sino una consecuencia de ella.

**30.** Sabemos que

$$\operatorname{sen} x = 2 \operatorname{sen}(x/2) \cos(x/2).$$

¿Está este hecho de acuerdo con $E$?

**31.** Predecir desde $E$ y verificar el valor del producto infinito

$$\left(1 - \frac{4}{1}\right)\left(1 - \frac{4}{9}\right)\left(1 - \frac{4}{25}\right)\left(1 - \frac{4}{49}\right) \cdots$$

**32.** Predecir desde $E$ y verificar el valor del producto infinito

$$\left(1 - \frac{16}{1}\right)\left(1 - \frac{16}{9}\right)\left(1 - \frac{16}{25}\right)\left(1 - \frac{16}{49}\right) \cdots$$

**33.** Comparar los ejemplos 31 y 32, y generalizar.

**34.** Sabemos que $\cos(-x) = \cos x$. ¿Está este hecho de acuerdo con $E$?

**35.** Sabemos que $\cos(x+\pi) = -\cos x$. ¿Está este hecho de acuerdo con $E$?

**36.** Derivar desde $E$ el producto para $1 - \operatorname{sen} x$ conjeturado en la sección 6 (4).

**37.** Derivar desde $E$ que

$$\cot x = \ldots + \frac{1}{x+2\pi} + \frac{1}{x+\pi} + \frac{1}{x} + \frac{1}{x-\pi} + \frac{1}{x-2\pi} + \ldots$$

**38.** Derivar desde $E$ que

$$\begin{aligned}\cot x = \frac{1}{x} &- \frac{2x}{\pi^2}\left(1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \ldots\right)\\ &- \frac{2x^3}{\pi^4}\left(1 + \frac{1}{16} + \frac{1}{81} + \frac{1}{256} + \frac{1}{625} + \ldots\right)\\ &- \frac{2x^5}{\pi^6}\left(1 + \frac{1}{64} + \frac{1}{729} + \ldots\right)\\ &- \ldots\end{aligned}$$

y encontrar la suma de la serie infinita cuyos coeficientes aparecen a la derecha.

**39.** Derivar desde $E$ que

$$\begin{aligned}\frac{\cos x}{1 - \operatorname{sen} x} &= \cot\left(\frac{\pi}{4} - \frac{x}{2}\right)\\ &= -2\left(\frac{1}{x-\frac{\pi}{2}} + \frac{1}{x+\frac{3\pi}{2}} + \frac{1}{x-\frac{5\pi}{2}} + \frac{1}{x+\frac{7\pi}{2}} + \ldots\right)\\ &= \frac{4}{\pi}\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \ldots\right)\\ &+ \frac{8x}{\pi^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49} + \frac{1}{81} \ldots\right)\\ &+ \frac{16x^2}{\pi^3}\left(1 - \frac{1}{27} + \frac{1}{125} - \frac{1}{343} + \ldots\right)\\ &+ \frac{32x^3}{\pi^4}\left(1 + \frac{1}{81} + \frac{1}{625} + \ldots\right)\\ &+ \ldots\end{aligned}$$

y encontrar la suma de la serie infinita cuyos coeficientes aparecen en la última expresión.

**40.** Mostrar que

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\frac{1}{25}+\ldots=\frac{4}{3}\left(1+\frac{1}{9}+\frac{1}{25}+\frac{1}{49}+\ldots\right)$$

que produce una segunda derivación para la serie de la izquierda.

**41.** (Continuación.) Intentar hallar una tercera derivación, sabiendo que

$$\text{arc sen } x = x+\frac{1}{2}\frac{x^3}{3}+\frac{1}{2}\frac{3}{4}\frac{x^5}{5}+\frac{1}{2}\frac{3}{4}\frac{5}{6}\frac{x^7}{7}+\ldots$$

y que para $n = 0, 1, 2, \ldots,$

$$\int_0^1 (1-x^2)^{-1/2}x^{2n+1}\,dx = \int_0^{\pi/2} (\text{sen } t)^{2n+1}\,dt = \frac{2\cdot 4\cdots 2n}{3\cdot 5\cdots(2n+1)}$$

**42.** (Continuación.) Intentar una cuarta derivación, sabiendo que

$$(\text{arc sen } x)^2 = x^2+\frac{2}{3}\frac{x^4}{2}+\frac{2}{3}\frac{4}{5}\frac{x^6}{3}+\frac{2}{3}\frac{4}{5}\frac{6}{7}\frac{x^8}{4}+\ldots$$

y que para $n = 0, 1, 2, \ldots,$

$$\int_0^1 (1-x^2)^{-1/2}x^{2n}\,dx = \int_0^{\pi/2} (\text{sen } t)^{2n}\,dt = \frac{1}{2}\frac{3}{4}\cdots\frac{2n-1}{2n}\frac{\pi}{2}.$$

**43.** Euler (*Opera Omnia*, ser. 1, vol. 14, págs. 40-41) usa la fórmula

$$1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\ldots$$

$$=\log x\cdot\log(1-x)+\frac{x+(1-x)}{1}+\frac{x^2+(1-x)^2}{4}+$$

$$+\frac{x^3+(1-x)^3}{9}+\ldots$$

válida para $0 < x < 1$. Calcular numéricamente la suma de la serie de la izquierda.

(*a*) Demostrar la fórmula.

(*b*) ¿Qué valor de $x$ es el más ventajoso para calcular la suma de la izquierda?

**44.** *Una objeción y una primera aproximación a la prueba.* No hay razón para admitir *a priori* que sen $x$ puede ser descompuesto en factores lineales correspondientes a las raíces de la ecuación

$$\operatorname{sen} x = 0.$$

Sin embargo, si admitiéramos esto aún habría una objeción: Euler *no* demostró que

$$0, \quad \pi, \quad -\pi, \quad 2\pi, \quad -2\pi, \quad 3\pi, \quad -3\pi, \quad \ldots$$

sean *todas* raíces de esta ecuación. A nosotros puede satisfacernos (analizando la curva $y = \operatorname{sen} x$) que no haya otras raíces reales, aunque Euler no excluyó de ningún modo la existencia de raíces complejas.

Esta objeción la levantó Daniel Bernouille (un hijo de Jean, 1700-1788). Euler contestó considerando que

$$\operatorname{sen} x = (e^{ix} - e^{-ix})/(2i)$$

$$= \lim_{n \to \infty} P_n(x),$$

donde

$$P_n(x) = \frac{1}{2i}\left[\left(1 + \frac{ix}{n}\right)^n - \left(1 - \frac{ix}{n}\right)^n\right]$$

es un polinomio (de $n$ grados, siendo $n$ impar).

Mostrar que $P_n(x)$ no tiene raíces complejas.

**45.** *Una segunda aproximación a la prueba.* Asumiendo que $n$ es impar en el ejemplo 44, factorizar $P_n(x)/x$ de modo que su factor $k$ se aproxime a

$$1 - \frac{x^2}{k^2\pi^2}$$

cuando $n$ tiende a $\infty$, para cualquier valor fijo de $k$ ($k = 1, 2, 3, \ldots$).

**46.** *Peligros de la analogía.* En resumen, la analogía entre lo finito y lo infinito condujo a Euler a un gran descubrimiento. Sin embago, dejó a un lado una falacia. Aquí tenemos un ejemplo que muestra el peligro a pequeña escala.

La serie

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \ldots = l$$

converge. Su suma $l$ puede ser groseramente estimada por los dos primeros términos

$$1/2 < l < 1;$$

ahora,

$$2l = \frac{2}{1} - \frac{1}{1} + \frac{2}{3} - \frac{1}{2} + \frac{2}{5} - \frac{1}{3} + \frac{2}{7} - \frac{1}{4} + \ldots$$

En esta serie hay sólo un término con un denominador par determinado (lo cual es negativo), pero dos términos con un denominador impar (un positivo y otro negativo). Juntemos los términos con el mismo denominador impar:

$$\begin{aligned} & \frac{2}{1} - \frac{1}{2} + \frac{2}{3} - \frac{1}{4} + \frac{2}{5} - \ldots \\ & - \frac{1}{1} \qquad - \frac{1}{3} \qquad - \frac{1}{5} \\ & = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots \\ & = l. \end{aligned}$$

Sin embargo, $2l \neq l$, puesto que $l \neq 0$. ¿Dónde está el error y cómo podemos evitar el repetirlo?

# Inducción en geometría sólida

*Aun en las ciencias matemáticas nuestros instrumentos principales para descubrir la verdad son la inducción y la analogía.*—LAPLACE *.

## 1. POLIEDROS

«Un complicado poliedro tiene muchas caras, vértices y aristas.» Alguna vaga observación de esta clase viene con facilidad a la mente del que haya tenido el menor contacto con geometría sólida. Sin embargo, no habrá muchas personas capaces de hacer un serio esfuerzo para profundizar en esta observación y buscar algo más preciso. Lo correcto sería distinguir claramente las cantidades envueltas y preguntar algo definido. En consecuencia, denotemos el número de caras, vértices y aristas del poliedro por $C$, $V$ y $A$, respectivamente (iniciales correspondientes a cada cosa), y preguntémonos esta cuestión: «¿Es cierto que incrementa el número de caras cuando incrementa el número de vértices? ¿Se incrementa $C$ necesariamente con $V$?».

Para empezar con poliedros particulares apenas podemos hacer nada mejor que examinar algunos ejemplos. Así, un cubo (el primer cuerpo, I, de la figura 3.1),

$$C = 6, \quad V = 8, \quad A = 12.$$

Ahora bien, para un prisma con base triangular (el cuerpo II en la figura 3.1),

$$C = 5, \quad V = 6, \quad A = 9.$$

Una vez en esta dirección, examinaremos y compararemos varios cuerpos, por ejemplo, los exhibidos en la figura 3.1, donde, junto a I y II, ya mencionados, tenemos también los siguientes: un prisma con base pentagonal (III), pirámides con base cuadrada, triangular o pentagonal (IV, V, VI), un octaedro (VII), una «torre con tejado» (VIII, una pirámide colocada sobre la cara superior de un cubo) y un «cubo truncado» (IX). Hagamos un pequeño esfuerzo de imaginación y represen-

* «Essai philosophique sur les probabilités», *Oeuvres Complètes de Laplace*, volumen 7, pág. v.

témonos estos cuerpos lo bastante claramente para contar las caras, vértices y aristas. Los números encontrados pueden verse en la siguiente tabla:

| | *Poliedros* | *C* | *V* | *A* |
|---|---|---|---|---|
| I. | Cubo | 6 | 8 | 12 |
| II. | Prisma triangular | 5 | 6 | 9 |
| III. | Prisma pentagonal | 7 | 10 | 15 |
| IV. | Pirámide cuadrada | 5 | 5 | 8 |
| V. | Pirámide triangular | 4 | 4 | 6 |
| VI. | Pirámide pentagonal | 6 | 6 | 10 |
| VII. | Octaedro | 8 | 6 | 12 |
| VIII. | «Torre» | 9 | 9 | 16 |
| IX. | «Cubo truncado» | 7 | 10 | 15 |

Nuestra figura 3.1 tiene alguna semejanza superficial con una representación mineralógica y la tabla citada es de alguna manera similar al libro de notas en que los científicos apuntan los resultados de sus descubrimientos. Examinemos y comparemos las figuras y los números de la tabla como el mineralogista y el físico examinan y comparan sus datos y modelos laboriosamente recogidos. Ahora, sin embargo, tenemos en nuestra mano algo con que contestar la pregunta: «¿Se incrementa *V* con *C*?». La contestación es «No»; comparando el cubo con el octaedro (I y VII) vemos que uno tiene más vértices y el otro más caras. Así, nuestro primer intento de establecer una completa regularidad ha fracasado.

Sin embargo, aún podemos intentar otra cosa. ¿Se incrementa *A* con *C*? ¿O con *V*? Para contestar a estas preguntas rehagamos nuestra tabla, disponiendo los poliedros de modo que *C* incremente cuando leamos los puntos sucesivos:

| *Poliedros* | *C* | *V* | *A* |
|---|---|---|---|
| Pirámide triangular | 4 | 4 | 6 |
| Pirámide cuadrada | 5 | 5 | 8 |
| Prisma triangular | 5 | 6 | 9 |
| Pirámide pentagonal | 6 | 6 | 10 |
| Cubo | 6 | 8 | 12 |
| Octaedro | 8 | 6 | 12 |
| Prisma pentagonal | 7 | 10 | 15 |
| «Cubo truncado» | 7 | 10 | 15 |
| «Torre» | 9 | 9 | 16 |

Al mirar ahora nuestros datos convenientemente ordenados podemos observar fácilmente que ninguna regularidad de la clase sospechada existe. Cuando $A$ incrementa de 15 a 16, $V$ cae de 10 a 9. De nuevo, cuando pasamos del octaedro al prisma pentagonal, $A$ incrementa de 12 a 15, pero $C$ cae de 8 a 7. Ni $C$ ni $V$ incrementan continuamente con $A$.

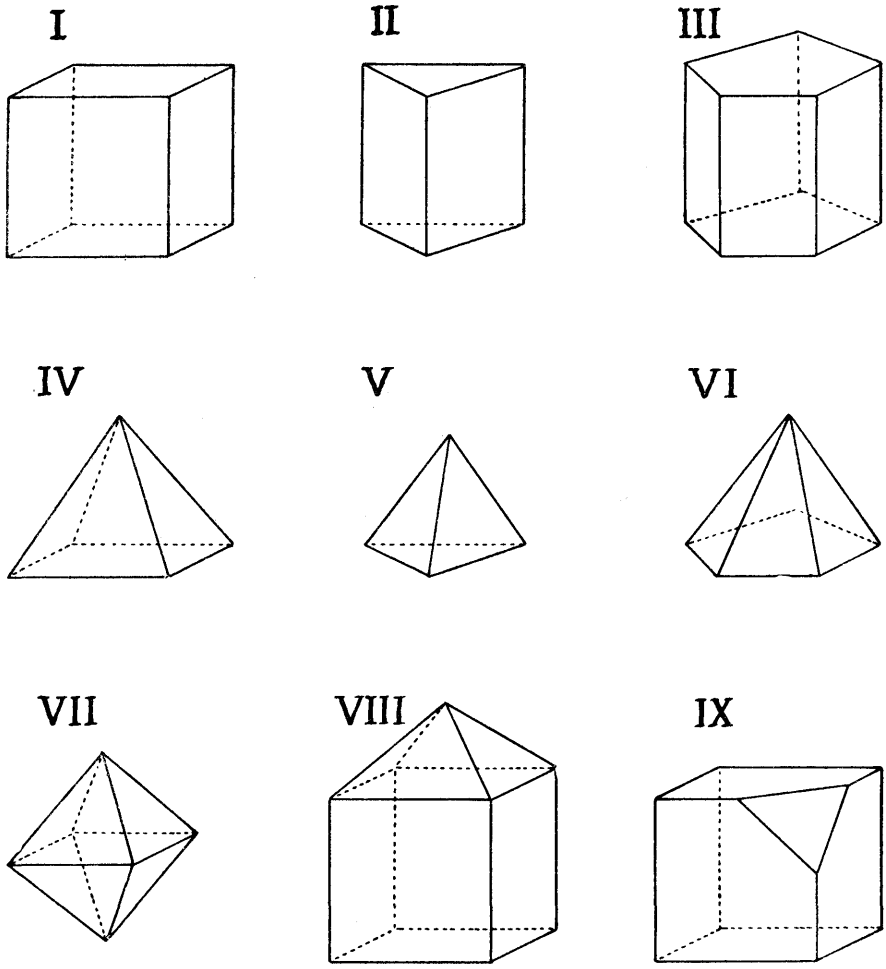



Fig. 3.1. Poliedros

Hemos fracasado otra vez en encontrar una regularidad válida. Sin embargo, no nos agrada admitir que nuestra idea original estaba completamente equivocada. Una cierta modificación de nuestra idea puede ser aún correcta. Ni $C$ ni $V$ incrementan con $A$, es verdad, pero parecen incrementar «en el conjunto». Examinando nuestros datos bien ajus-

tados podemos observar que $C$ y $V$ incrementan «conjuntamente»: $C + V$ *incrementan* como se lee abajo. Y luego una regularidad más precisa nos sorprende a lo largo de la tabla

$$C + V = A + 2.$$

Esta relación se verifica en los nueve casos de la tabla. No parece que una tan persistente regularidad sea mera coincidencia. Así, nos vemos conducidos a la conjetura de que no sólo en los casos observados, sino de que en cualquier poliedro *el número de caras más el número de vértices es igual al número de aristas más dos.*

## 2. Primeros contactos «de apoyo»

Un naturalista riguroso no admite fácilmente una conjetura. Aunque la conjetura se presente como plausible y haya sido verificada en algunos casos, se la planteará y recogerá nuevas observaciones o diseñará nuevos experimentos para probarla. Nosotros podemos hacer la misma cosa. Vamos a examinar aún otros poliedros, a contar sus caras, vértices y aristas, y a comparar $C + V$ con $A + 2$. Estos números pueden ser iguales o no. Será interesante ver lo que resulta.

Mirando a la figura 3.1 podemos observar que ya hemos examinado tres cuerpos regulares: el cubo, el tetraedro y el octadero (I, V, VII). Vamos a examinar los dos que quedan: el icosaedro y el dodecaedro.

El icosaedro tiene veinte caras, todas triángulos, y así, $C = 20$. Los veinte triángulos tienen en total $3 \times 20 = 60$ lados, de los cuales dos coinciden en la misma arista del icosaedro. Por lo tanto, el número de aristas es $60/2 = 30 = A$. Podemos encontrar análogamente $V$ como sigue. Sabemos que cinco caras del icosaedro se agrupan cada una alrededor de cadavértice. Los veinte triángulos tienen en total $3 \times 20 = 60$ ángulos, de los cuales cinco pertenecen al mismo vértice. Por lo tanto el número de vértices es $60/5 = 12 = V$.

El dodecaedro tiene doce caras, todas ellas pentágonos, de las cuáles tres se agrupan alrededor de cada vértice. De aquí podemos concluir, semejantemente a los casos anteriores, que

$$C = 12, \quad V = \frac{12 \times 5}{3} = 20, \qquad A = \frac{12 \times 5}{2} = 30.$$

Ahora podemos añadir a nuestra tabla dos líneas más:

| *Poliedros* | $C$ | $V$ | $A$ |
|---|---|---|---|
| Icosaedro . . . . . | 20 | 12 | 30 |
| Dodecaedro . . . . . | 12 | 20 | 30 |

Nuestra conjetura de que $C + V = A + 2$ se verifica en ambos casos.

## 3. Más contactos «de apoyo»

Gracias a las precedentes verificaciones nuestra conjetura se ha hecho sensiblemente más plausible, pero ¿está probada ya? De ninguna manera. En una situación semejante un naturalista concienzudo sentiría satisfacción por el éxito de sus experimentos, pero continuaría discurriendo otros nuevos. ¿Qué poliedro probaríamos ahora?

La cuestión es que nuestra conjetura ha sido tan bien verificada hasta ahora que verificarla una vez más añadiría muy poco a nuestra confianza, tan poco quizá que escasamente merecería la pena la molestia de elegir un poliedro y contar sus partes. ¿Podríamos encontrar alguna manera más valiosa para probar nuestra conjetura?

Mirando a la figura 3.1 podemos observar que los cuerpos sólidos de la primera línea son todos de la misma naturaleza: prismas. De nuevo vemos que todos los de la segunda línea son pirámides. Ciertamente, nuestra conjetura es verdadera en los tres prismas y las tres pirámides mostrados en dicha figura; pero ¿es verdadera de *todos* los prismas y pirámides?

Si un prisma tiene $n$ caras laterales, tiene $n + 2$ caras en total, $2n$ vértices y $3n$ aristas. Una pirámide con $n$ caras laterales tiene $n + 1$ caras en total, $n + 1$ vértices y $2n$ aristas. De esta manera, podemos añadir dos líneas más a nuestra tabla:

| *Poliedros* | $C$ | $V$ | $A$ |
|---|---|---|---|
| Prisma con $n$ caras laterales . . . . | $n + 2$ | $2n$ | $3n$ |
| Pirámide con $n$ caras laterales . . . | $n + 1$ | $n + 1$ | $2n$ |

Nuestra conjetura afirmando que $C + V = A + 2$ resulta verdadera no sólo para uno o dos poliedro más, sino para dos series ilimitadas de poliedros.

## 4. Una prueba rigurosa

La última observación aumenta considerablemente nuestra confianza en la conjetura, pero, naturalmente, no la prueba. ¿Qué haríamos? ¿Seguiríamos probando nuevos casos particulares? Nuestra conjetura parece salir favorecida de las pruebas simples. Por tanto, sometámosla a alguna más rigurosa que suponga una oportunidad de rechazarla.

Vamos a mirar de nuevo a nuestra colección de poliedros (figura 3.1). Tenemos prismas (I, II, III), pirámides (IV, V, VI), cuerpos regulares (I, V, VII); sin embargo, ya hemos considerado todas estas clases de cuerpos exhaustivamente. ¿Qué otra cosa queda? La figura 3.1 contiene también la «torre» (VIII), obtenida colocando un «tejado» sobre un cubo. Aquí podemos percibir la posibilidad de una generalización. Tomamos cualquier poliedro en vez del cubo, elegimos una cara del poliedro y colocamos un «tejado» sobre ella. Supongamos que el poliedro original tiene $C$ caras, $V$ vértices y $A$ aristas, y supongamos que la cara elegida tiene $n$ lados. Sobre esta cara colocamos una pirámide con $n$ caras laterales y así obtenemos un nuevo poliedro. ¿Cuántas caras, vértices y aristas tiene el nuevo poliedro «techado»? Una cara (la cara elegida) se pierde en el proceso y se ganan $n$ nuevas (las $n$ caras laterales de la pirámide), de manera que el nuevo poliedro tiene $C - 1 + n$ caras. Todos los vértices del poliedro pertenecen también al nuevo, y aún uno más (la cúspide de la pirámide), de modo que el nuevo poliedro tiene $V + 1$ vértices. Por otro lado, todas las aristas del antiguo poliedro las tiene también el nuevo y aún se añaden $n$ aristas (las laterales de la pirámide), de modo que el nuevo poliedro tiene $A + n$ aristas.

Resumamos: el poliedro original tenía $C$, $V$ y $A$ caras, vértices y aristas, respectivamente, mientras el nuevo poliedro «techado» tiene

$$C + n - 1, \quad V + 1 \quad \text{y} \quad A + n$$

partes de la clase correspondiente. ¿Están estos hechos de acuerdo con nuestra conjetura?

*Si* la relación $C + V = A + 2$ se mantiene, entonces, obviamente,

$$(C + n - 1) + (V + 1) = (A + n) + 2$$

se mantiene *también*. Es decir, si nuestra conjetura se verifica en el caso del poliedro original, también debe verificarse en el caso del nuevo

poliedro «techado». Nuestra conjetura supera el «techado», y, así, ha pasado una prueba realmente rigurosa. Hay una variedad inagotable de poliedros que podemos derivar de los ya examinados, «techándolos» también, y probar así que nuestra conjetura es cierta para ellos.

A propósito, el último cuerpo de nuestra figura 3.1, el «cubo truncado» (IX), abre la vía a una consideración similar. En lugar del cubo cojamos cualquier poliedro «truncado», cortando arbitrariamente el vértice elegido. Sabemos que el poliedro original tiene

$$C, V \text{ y } A$$

caras, vértices y aristas, respectivamente, y sabemos que $n$ es el número de aristas que salen del vértice elegido. Cortemos este vértice e introduzcamos una nueva cara (que tiene $n$ lados), $n$ nuevas aristas, $n$ nuevos vértices, aunque ha perdido uno de éstos también. La suma del nuevo poliedro «truncado» será

$$C + 1, \quad V + n - 1 \quad \text{y} \quad A + n$$

caras, vértices y aristas, respectivamente. Ahora, de

$$C + V = A + 2$$

se sigue

$$(C + 1) + (V + n - 1) = (A + n) + 2.$$

Es decir, nuestra conjetura es lo bastante tenaz para superar el «truncamiento». Ha pasado otra prueba rigurosa.

Es natural considerar las anteriores observaciones como un argumento muy fuerte en favor de nuestra conjetura. Podemos percibir en ellas algo más: el primer índice de una prueba. Empezando con un poliedro simple, como el tetraedro o el cubo, para los que la conjetura es válida, podemos pasar, «techando» y «truncando», a una gran variedad de poliedros para los que la conjetura también es válida. ¿Podríamos pasar a ***todos*** los poliedros? ¡Eso sería tener una prueba! Además, hay otras operaciones que, igual que las anteriores, mantienen la conjetura.

## 5. Verificaciones y verificaciones

Los procedimientos de un naturalista concienzudo no son esencialmente diferentes de los que emplea el hombre común, aunque

sí más eficaces. El hombre común y el científico llegan a conjeturas por unas pocas observaciones y prestan atención a los últimos casos que concuerdan o no con la conjetura. Un caso de concordancia da a la conjetura más crédito, un caso en desacuerdo la desaprueba totalmente, y aquí empiezan las diferencias: las personas corrientes son más aptas para buscar la primera clase de casos, mientras los científicos son más aptos para los segundos. La razón es que todos tenemos algo de vanidad, tanto el hombre de la calle como el científico, pero cada uno pone el orgullo en cosas diferentes. A don Fulano de Tal no le gusta confesar, aunque sea a sí mismo, que está equivocado y por ello no le gustan los casos en contra; por el contrario, los evita y se siente inclinado a minosvalorarlos cuando se presentan. El científico, al revés, está preparado para reconocer las conjeturas equivocadas, pero no le gusta abandonar los problemas sin resolver. Ahora bien, un caso de concordancia no asegura nada definitivamente, aunque un caso en contra lo haga. El científico, buscando una decisión definitiva, procura casos que desbaraten la conjetura, y cuanto más oportunidad de ello ofrecen mejor acogidos son. Hay aquí un punto importante que observar. Si un caso que amenazaba desbaratar la conjetura es rechazado resultando al final estar de acuerdo con ella, la conjetura sale grandemente fortalecida de la prueba. A mayor peligro mayor honor; pasar los exámenes más duros garantiza el reconocimiento más alto, la mayor evidencia experimental de la conjetura. Hay ejemplos y ejemplos, verificaciones y verificaciones. Un ejemplo que está más en pugna con la conjetura la pone en cualquier caso más cerca de la decisión que un ejemplo menos en pugna con ella, y esto explica la preferencia del científico.

Ahora podemos pasar a nuestro problema particular y ver cómo las observaciones precedentes pueden ser aplicadas a la «investigación experimental sobre poliedros» que hemos emprendido. Cada nuevo caso en que la relación $C + V = A + 2$ se verifica añade confianza a su verdad generalmente admitida. Sin embargo, pronto nos cansaremos de una monótona secuencia de verificaciones. Un caso algo diferente de los examinados añade confianza, si está de acuerdo con nuestra conjetura, pero no demasiada. De hecho, antes de la prueba podemos ya suponer que el caso entre manos se conducirá como el anterior, puesto que difiere tan poco de él. Nuestro deseo, pues, no es sólo una verificación más, sino una *verificación de otra clase.* Y, en realidad, si volvemos la vista a las varias fases de nuestra investigación (secciones 2, 3, 4) observaremos que cada verificación supera las obtenidas anteriormente. Cada fase verifica la conjetura para una variedad de casos más extensa que la precedente.

## 6 Un caso muy diferente

Siendo la variedad tan importante, busquemos algunos casos muy diferentes de los examinados hasta aquí. Podemos, por ejemplo, considerar el marco de un cuadro como un poliedro. Cojamos un palo largo y triangular, cortémosle en cuatro partes, peguemos los bordes cortados y nos encontraremos con un poliedro en forma de marco. La figura 3.2

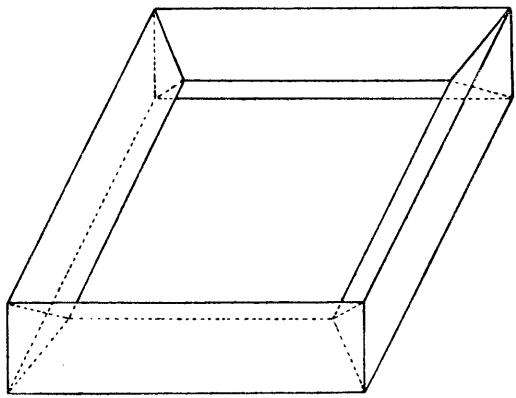


Fig. 3.2. Un poliedro con «forma de buñuelo».

sugiere que el marco está colocado sobre una mesa de modo que las aristas no cortadas yacen horizontalmente. Hay cuatro veces tres, es decir, doce aristas horizontales, y también cuatro veces tres aristas no horizontales, de modo que el número total de aristas es $A = 12 + 12 = 24$. Si contamos caras y vértices tenemos que $C = 4 \times 3 = 12$ y $V = 4 \times 3 = 12$. Ahora bien, $C + V = 24$ es diferente de $A + 2 = 26$. Nuestra conjetura, tomada con completa generalidad, ¡se ha vuelto falsa!

Sin duda, podemos decir que nunca hemos intentado establecer la conjetura con tanta generalidad, que nosotros nos referíamos a poliedros convexos o, digamos, con «forma esférica», y no a poliedros con «forma de buñuelo» como el marco de un cuadro. Pero todo ello son excusas. De hecho, nosotros hemos cambiado nuestra posición y modificado el juicio original. Es bastante posible que el choque que hemos sufrido sea beneficioso en definitiva y nos conduzca a enmendar y establecer un enunciado más preciso de nuestra conjetura. Sin embargo, y en cualquier caso, ha sido un golpe para nuestra confianza.

## 7. Analogía

El ejemplo del «marco del cuadro» destruyó nuestra conjetura en su forma original, pero pronto ha resucitado bajo una forma revisada (y esperemos que mejorada), con una importante restricción.

El tetraedro es convexo, y lo mismo el cubo y los otros poliedros de nuestra colección (figura 3.1), igual que los poliedros «techados» y «truncados» que derivamos de aquéllos. En cualquier caso, no hay peligro de que estas operaciones nos condujeran de un poliedro convexo o de «forma esférica» a un cuerpo sólido de «forma de buñuelo».

Al observar esto es conveniente introducir alguna mayor precisión. Habíamos conjeturado que en cualquier poliedro *convexo* se da la relación $C + V = A + 2$ entre caras, vértices y aristas. (Es preferible restringirlos a poliedros de «forma esférica», pero ahora no deseo pararme a definir el significado del término.)

Esta conjetura tiene alguna probabilidad de ser cierta. Sin embargo, nuestra confianza se ha derrumbado y necesitamos algún nuevo apoyo para nuestra conjetura. No podemos esperar demasiado de más verificaciones, pues parece que hemos agotado las fuentes más obvias. Aun así, todavía podemos esperar algo de la analogía. ¿Hay algún caso análogo que nos pueda ser provechoso?

Los polígonos son análogos a los poliedros. Un polígono es parte de una superficie como un poliedro es parte del espacio. Un polígono tiene cierto número de vértices, $V$ (los vértices de sus ángulos), y un cierto número, $A$, de aristas (las de sus lados). Es obvio, pues, que

$$V = A.$$

Sin embargo, esta relación, válida para polígonos convexos, resulta demasiado simple y arroja poca luz sobre la relación más intrincada

$$C + V = A + 2,$$

que sospechamos es válida para los poliedros convexos.

Si estamos de veras interesados en el problema trataremos de acercar estas relaciones una a otra. Hay una ingeniosa manera de hacerlo: procuraremos poner los distintos números dentro de un orden natural. El poliedro es 3-dimensional; sus caras (polígonos) son 2-dimensionales, sus aristas 1-dimensionales y sus vértices 0-dimensionales, lógicamente. Ahora volvamos a escribir las ecuaciones en el orden del incre-

mento de las dimensiones. La relación para los polígonos, escrita así:

$$V - A + 1 = 1,$$

es comparable a la relación para poliedros, escrita así:

$$V - A + C - 1 = 1.$$

El 1 en el lado izquierdo de la ecuación para polígonos vale por el concerniente elemento bidimensional, el interior del polígono. El 1 en el lado izquierdo de la ecuación para poliedros vale para el concerniente elemento tridimensional, el interior del poliedro. Los elementos del lado izquierdo cuentan 0, 1, 2 y 3 dimensiones, respectivamente; están dispuestos en este orden natural y tienen signos alternativos. El lado derecho es el mismo en ambas ecuaciones; la analogía parece completa. Puesto que la primera ecuación para polígonos es obviamente cierta, la analogía añade confianza a la segunda ecuación conjeturada para poliedros.

## 8. La división del espacio

Ahora pasamos a otro ejemplo de investigación inductiva en geometría sólida. En el anterior ejemplo empezamos con una observación general y bastante vaga. Nuestro punto de partida es ahora un problema particular muy claro. Se trata de un problema sencillo de geometría sólida, aunque no muy familiar: *¿En cuántas partes queda dividido el espacio por cinco planos?*

Este problema es fácil de contestar si los cinco planos son todos paralelos, en cuyo caso el espacio queda dividido en seis partes. Este caso es, sin embargo, demasiado particular. Si los planos están en una «posición general», de modo que entre ellos no haya dos que sean paralelos, entonces tendremos bastantes más de seis partes. Podemos ahora volver a formular el problema añadiéndole una cláusula esencial: *¿En cuántas partes es dividido el espacio por cinco planos, habida cuenta de que estos planos están en una posición general?*

La idea de una «posición general» es bastante intuitiva; queremos decir que los planos están en una posición tal que no se hallan ligados entre sí por una relación particular, es decir, cuando se dan independientemente, elegidos al azar. No sería difícil aclarar el término con una definición técnica, pero no lo hacemos así por varias razones. Primera,

porque dicha presentación no sería demasiado técnica. Segunda, al dejar la noción de algún modo nebulosa estamos más cerca de la actitud mental del naturalista, obligado a menudo a empezar con nociones nebulosas, aclarándolas después sobre la marcha.

## 9. Modificando el problema

Concentrémonos sobre nuestro problema. Tenemos cinco planos en posición general, que cortan el espacio en un cierto número de divisiones. (Podemos pensar, por ejemplo, en las rodajas de un queso dividido en trozos por cinco cortes de un afilado cuchillo.) El problema es, entonces, encontrar el número de divisiones. (¿En cuántos trozos quedó cortado el queso?)

Parece difícil ver en seguida las divisiones efectuadas por los cinco planos. (Puede ser imposible *verlas*. En todo caso, no debemos esforzar nuestra imaginación geométrica; por el contrario, intentemos pensar. Su razón puede llevarle más lejos que su imaginación.) Pero, ¿por qué sólo cinco planos? ¿Por qué no cualquier número de planos? ¿En cuántas partes es dividido el espacio por cuatro planos? ¿Por tres planos? ¿O por dos? ¿O sólo por uno?

Nosotros visualizamos así casos que son accesibles a la intuición geométrica. Un plano divide el espacio obviamente en dos partes. Dos planos dividen el espacio en tres partes, siempre que sean paralelos. Pero antes hemos descartado las posiciones particulares; dos planos en una posición general interceden y dividen el espacio en cuatro partes; tres planos en posición general lo dividen en ocho. Con objeto de comprobar este último y más difícil caso podemos pensar en dos paredes verticales dentro de un edificio, cruzándose una a otra y en una capa horizontal, soportada por vigas, cruzando ambas paredes y formando alrededor del punto donde se cruzan ambas el techo de cuatro habitaciones y el suelo de otras cuatro.

## 10. Generalización, especialización, analogía

Nuestro problema tiene que ver con cinco planos, pero antes de considerar cinco planos hemos trabajado con uno, dos y tres planos. ¿Hemos malgastado nuestro tiempo? De ninguna manera; al contrario, nos hemos facilitado el problema examinando *casos análogos más sencillos*. Nos hemos familiarizado con estos casos más sencillos, hemos

aclarado los conceptos que intervienen y hemos allanado el camino para resolver el problema con que tenemos que enfrentarnos.

Pero este mismo camino que nos conduce a problemas análogos más simples es típico y merece él mismo nuestra atención. Primero pasamos del caso de cinco planos al de cualquier número de planos, digamos, a $n$ planos: hemos *generalizado*. Luego, de $n$ planos hemos vuelto a cuatro, a tres, a dos planos, incluso a un solo plano, es decir, pusimos $n = 4, 3, 2, 1$ en el problema general: hemos *especializado*. Pero el problema de la división del espacio en tres planos es *análogo*, digamos, a nuestro problema inicial, que envuelve cinco planos. Así, hemos practicado la analogía de una manera típica, por *generalización introducida y especialización subsecuente*.

## 11. Un problema análogo

¿Qué decir sobre el próximo caso de cuatro planos?

Cuatro planos, en posición general, determinan varias porciones de espacio, una de las cuales es límite; contiene cuatro caras triangulares y es llamado tetraedro (ver figura 3.3). Esta configuración nos

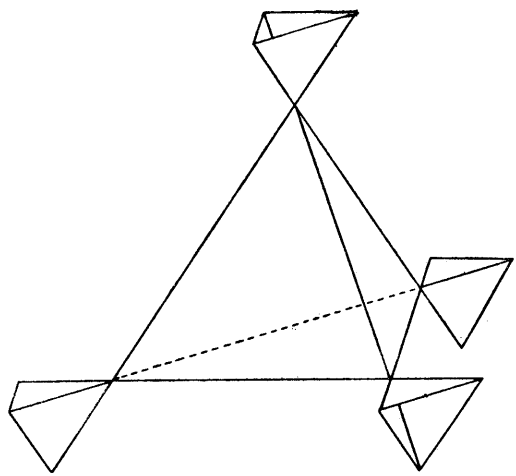


Fig. 3.3. Espacio dividido por cuatro planos.

recuerda la de tres líneas rectas en un plano, en posición general, que determina varias porciones del plano, una de las cuales es límite y contiene tres segmentos línea, recibiendo el nombre de triángulo (ver figura 3.4). Ahora debemos averiguar el número de porciones de espacio determinadas por cuatro planos. Intentemos hacernos con el problema

análogo más simple: ¿En cuántas porciones es dividido el espacio por tres líneas?

Muchos de nosotros vemos la contestación inmediatamente aun sin dibujar una figura, pero, en todo caso, cualquiera puede verlo usando un grosero esquema (ver figura 3.4). El número requerido de partes es siete.

Hemos encontrado así la solución del problema análogo más sencillo. Pero, ¿nos servirá esta solución para nuestro problema inicial?

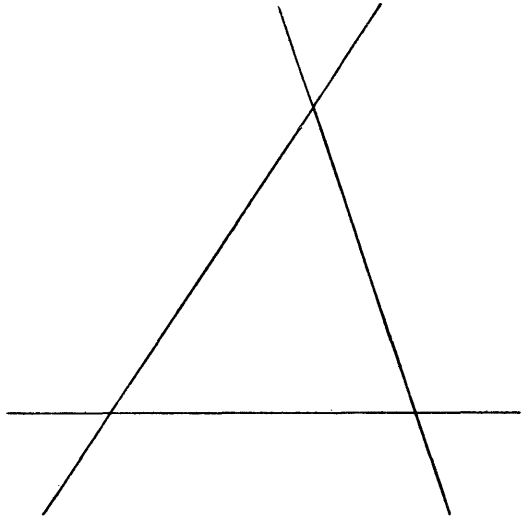


Fig. 3.4. Superficie dividida por tres líneas.

Sí, desde luego, siempre que utilicemos inteligentemente la analogía de las dos configuraciones. Consideremos la disección del plano por tres rectas de modo que podamos después aplicar la misma consideración a la disección del espacio por cuatro planos.

Así, miremos otra vez la disección del plano por tres líneas, limitando un triángulo. Una división es finita, la del interior del triángulo. Y las divisiones infinitas tienen todas un lado común con el triángulo (hay tres divisiones tales), o un vértice común (hay también tres divisiones de esta clase). El número total de divisiones es, pues, $1 + 3 + 3 = 7$.

Ahora consideremos la disección del espacio por cuatro planos limitando un tetraedro. Una división es finita, la del interior del tetraedro. Una división infinita puede tener una cara común (una 2-dimensional parte del límite) con el tetraedro (hay cuatro divisiones tales), o una arista común (una parte 1-dimensional del límite; hay seis divi-

siones de esta clase), o un vértice común (una parte 0-dimensional del límite; hay cuatro divisiones de esta clase, enfatizadas en la figura 3.3). El número total de divisiones será, entonces, $1 + 4 + 6 + 4 = 15$.

Hemos alcanzado este resultado por analogía, utilizándola de una manera característica e importante. Primero, hemos proyectado un problema análogo más fácil y lo hemos resuelto. Luego, con el objeto de resolver el problema inicial más difícil (sobre el tetraedro), hemos utilizado el problema más fácil (sobre el triángulo) como *modelo;* al resolver el problema más difícil seguimos el patrón de solución para el más fácil. Pero, antes de hacer esto, tuvimos que reconsiderar la solución del primer problema y rehacerlo dentro de un patrón adecuado para ser imitado.

Separar un problema análogo más fácil, resolverlo, rehacer su solución de modo que sirva de modelo y, por fin, alcanzar la solución del problema inicial siguiendo el modelo creado es un método que puede parecer un rodeo a los no iniciados, pero frecuentemente usado en la investigación científica, tanto matemática como no matemática.

## 12. Una colección de problemas análogos

Sin embargo, nuestro problema inicial sigue todavía sin resolver. Se refiere a la disección del espacio por cinco planos. ¿Cuál es el problema análogo en dos dimensiones? ¿La disección por cinco líneas rectas? ¿O por cuatro? Puede ser mejor considerar estos problemas con completa generalidad, esto es, la disección del espacio por $n$ planos y la del plano por $n$ rectas. Estas líneas deben estar, como es lógico, en posición general (que no haya dos paralelas ni tres que se unan en el mismo punto).

Si estamos acostumbrados a utilizar la analogía geométrica podemos dar un paso más y considerar también la división de la línea recta por $n$ puntos diferentes. Aunque esto parece ser trivial, puede resultar instructivo. Podemos ver fácilmente que una línea recta queda dividida en dos partes por un punto, en tres partes por dos puntos, en cuatro partes por tres puntos, y, en general, en $n + 1$ partes por $n$ puntos.

Por otro lado, si estamos acostumbrados a poner atención en los casos extremos, podemos considerar la línea, la superficie y el plano como indivisos y mirarla como «una división efectuada por 0 elementos».

Establezcamos la siguiente tabla con todos los resultados obtenidos hasta aquí:

| Número de elementos que dividen | Número de divisiones | | |
|---|---|---|---|
| | Del espacio por planos | Del plano por líneas rectas | De la línea por puntos |
| 0 | 1 | 1 | 1 |
| 1 | 2 | 2 | 2 |
| 2 | 4 | 4 | 3 |
| 3 | 8 | 7 | 4 |
| 4 | 15 | | 5 |
| ... | ... | ... | ... |
| $n$ | | | $n+1$ |

## 13. Muchos problemas pueden ser más fáciles que uno solo

Nosotros empezamos a resolver un problema sobre la disección del espacio por cinco planos. Aún no lo hemos resuelto, aunque hemos propuesto muchos nuevos problemas. Cada caso de nuestra tabla aún no relleno corresponde a una cuestión abierta.

Este procedimiento de amontonar nuevos problemas puede parecer tonto al no iniciado. Pero la experiencia de resolver problemas nos enseña que muchos problemas juntos pueden ser más fáciles de resolver que uno solo de ellos —si los muchos problemas están coordinados y el otro se encuentra aislado—. Ahora, nuestro problema inicial aparece como uno más en una colección de problemas sin resolver. Pero lo importante es, precisamente, que todos estos problemas sin resolver estén en una colección, es decir: dispuestos, agrupados, en íntima analogía unos con otros y unos pocos resueltos ya. Si comparamos la situación actual de la cuestión, inserta en una colección de casos análogos, con la situación original, cuando estaba totalmente aislada, nos sentimos, naturalmente, inclinados a creer que algún progreso ha sido hecho.

## 14. Una conjetura

Miremos los resultados desplegados en nuestra tabla como un naturalista mira la colección de sus muestras. Esta tabla es un desafío a nuestra habilidad inventiva, a nuestras facultades de observación. ¿Podemos descubrir en ella alguna conexión, alguna regularidad?

Mirando la segunda columna (división del espacio por planos)

podemos observar la secuencia 1, 2, 4, 8 —es decir, una clara regularidad—; vemos en ella una sucesión de valores de 2. Sin embargo, ¡qué desilusión! El próximo término en la columna es 15 y no 16, como esperábamos. Nuestra primera intuición no ha sido muy buena; busquemos otra.

Eventualmente, podemos coger dos números yuxtapuestos y observar su suma en la tabla. En seguida reconocemos una peculiar conexión; nosotros obtenemos un número de la tabla añadiendo otros dos, el número de arriba y el último de la derecha. Por ejemplo,

$$\begin{matrix} 8 & & 7 \\ 15 & & \end{matrix}$$

son enlazados por la relación

$$8 + 7 = 15.$$

Esta es una observación digna de nota, una clave sorprendente. Pues no parece que esta conexión, observada a lo largo de toda la tabla hasta donde ha sido calculada, sea una mera coincidencia.

Así, la situación sugiere que la regularidad observada se extiende más allá de los límites de nuestra observación, que los números de la tabla todavía no calculados están conectados de la misma manera que los calculados, y esto nos conduce a la conjetura de que le ley encontrada es generalmente válida.

Y *si* esto es así, ya estamos en condiciones de resolver nuestro problema original. Por adición de números yuxtapuestos podemos extender la tabla hasta alcanzar el número que deseamos obtener:

| | | | |
|---|---|---|---|
| 0 | 1 | 1 | 1 |
| 1 | 2 | 2 | 2 |
| 2 | 4 | 4 | 3 |
| 3 | 8 | 7 | 4 |
| 4 | 15 | **11** | 5 |
| 5 | **26** | | |

En la tabla reimpresa aquí aparecen dos nuevos números en tipos gruesos, calculados por adición, $11 = 7 + 4$, $26 = 15 + 11$. Si nuestra intuición es correcta, veintiséis será el número de partes en que el espacio

es dividido por cinco planos en posición general. Parece, pues, que hemos resuelto el problema. O, al menos, que hemos conseguido atinar con una conjetura plausible y apoyada por todas las evidencias recogidas hasta aquí.

## 15. Predicción y verificación

En lo anterior hemos seguido exactamente el procedimiento típico del naturalista. Si un naturalista observa una regularidad chocante, que no puede ser razonablemente atribuida a mera oportunidad, conjetura que la regularidad se extiende más allá de los límites de sus observaciones actuales. Hacer tal conjetura es, a menudo, el paso decisivo en la investigación inductiva.

El paso siguiente puede ser dado por predicción. Sobre las bases de las primeras observaciones y su concordancia con la ley conjetural el naturalista predice el resultado de su próxima observación. La mayor parte depende del resultado de esa próxima observación. ¿Será la predicción cierta o no? Estamos en la misma posición. Hemos encontrado, o mejor, intuido o predicho que once será el número de porciones en que

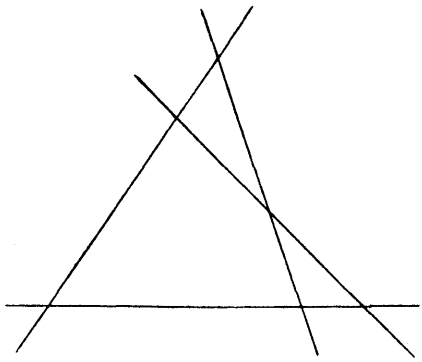


Fig. 3.5. Plano dividido por cuatro líneas.

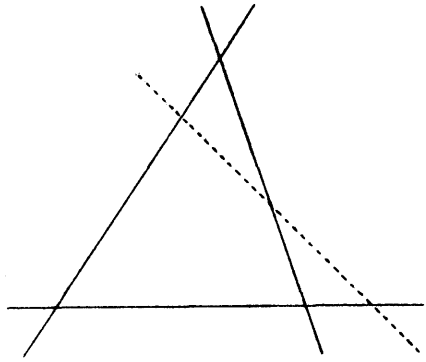


Fig. 3.6. Transición de tres líneas a cuatro.

un plano es dividido por cuatro rectas en posición general. ¿Es eso así? ¿Es nuestra predicción correcta?

Examinando un grosero esquema (ver figura 3.5) podemos convencernos de que nuestra intuición era buena, que once *es* realmente el número correcto. Esta confirmación de nuestra predicción conduce la evidencia inductiva en favor de la regla sobre cuya base hicimos nuestra

predicción. Habiendo pasado la prueba con éxito nuestra conjetura sale fortalecida.

## 16. Otra vez y mejor

Hemos verificado el número 11 mirando la figura y contando. Sí; cuatro líneas en posición general parecen dividir el plano en once porciones. Pero, no importa; hagámoslo otra vez y mejor. Hemos contado las porciones de alguna manera. Contémolas otra vez de manera que estemos seguros de evitar toda confusión y equivocación, o caída en posiciones especiales.

Empezamos por el hecho de que tres líneas determinan exactamente siete porciones del plano. Tenemos también razones para creer que cuatro líneas determinan once porciones. ¿Por qué sólo cuatro porciones más? ¿Por qué interviene el número 4 en esta conexión? ¿Por qué la introducción de una nueva línea incrementa el número de porciones sólo en cuatro?

En la línea de la figura 3.5 hemos enfatizado una línea. Hemos vuelto a dibujarla con trazos ligeros en la figura 3.6. La segunda figura no parece muy diferente, pero expresa una concepción diferente. Ahora enfatizamos dicha línea como *nueva*, mientras consideramos *viejas* a las otras tres. Las líneas viejas cortan la superficie en siete porciones. ¿Qué sucede cuando se añade una nueva línea?

La nueva línea, dibujada al azar, debe cortar cada una de las líneas viejas en un punto diferente. Así se constituyen tres puntos. Estos tres puntos dividen la nueva línea en cuatro segmentos. Ahora cada segmento secciona en dos una división del plano, haciendo dos nuevas divisiones de una vieja. Tomados juntos, los cuatro segmentos de la nueva línea crean ocho nuevas divisiones y abolen cuatro de las viejas —el *número* de divisiones incrementa, pues, en cuatro—. *Esta* es la razón por la que el número de divisiones es ahora cuatro más que antes: $7 + 4 = 11$.

Esta manera de llegar al número 11 es convincente e ilustrativa. Ahora empezamos a ver la razón de la regularidad que habíamos observado y sobre la que habíamos basado la predicción del número 11. Empezamos a sospechar una explicación por detrás de los hechos y nuestra confianza en la validez de la regularidad observada se vio muy fortalecida.

## 17. La inducción sugiere la deducción; el caso particular sugiere la prueba general

Nos hemos preocupado mucho de señalar el paralelismo entre nuestro razonamiento y los procedimientos del naturalista. Empezamos con un problema particular como el naturalista puede empezar con una observación enigmática. Luego fuimos progresando por generalizaciones tentativas, viendo casos particulares más accesibles y observando analogías instructivas. Más tarde intentamos intuir alguna regularidad y fracasamos, volvimos a intentarlo y nos fue mejor; logramos conjeturar una ley general que fue confirmada por toda la evidencia experimental a nuestra disposición; pusimos a prueba un caso particular más y encontramos que concordaba con la ley conjeturada, con cuya verificación ganó mucha autoridad. Al fin, empezamos a ver una razón de esa ley general, una especie de explicación, y nuestra confianza se vio muy fortalecida. La investigación de un naturalista habría pasado exactamente por las mismas fases.

Hay, sin embargo, un punto decisivo en que el matemático se separa del naturalista. La observación es la autoridad más alta para el naturalista, aunque no para el matemático. La verificación en muchos ejemplos bien elegidos es la única manera de confirmar una ley conjetural en ciencias naturales. Por el contrario, en matemáticas tales ejemplos bien elegidos pueden ser muy valiosos como estímulos, pero nunca bastan para probar una ley conjetural. Consideremos nuestro mismo caso concreto. Por el examen y comparación de varios casos particulares llegamos a conjeturar una regla general de la que se seguía que la solución a nuestro problema inicial era 26. ¿Son todas nuestras observaciones y verificaciones suficientes para probar la regla general? ¿O pueden probar que la solución de nuestro problema particular es realmente 26? No, en última instancia. Para un matemático riguroso el número 26 es sólo una aguda intuición y no hay cantidad de experiencia que pueda demostrar la ley general sospechada. La inducción convierte en probable el resultado, pero no lo llega a probar.

Podemos observar, sin embargo, que la investigación inductiva resulta útil en matemáticas en un aspecto que no hemos aún mencionado. La cuidadosa observación de casos particulares que conducen a un resultado matemático general puede también *sugerir su prueba*. Del intento de examen de un caso particular puede surgir un atisbo general.

De hecho, esto nos ha sucedido ya en la anterior sección. La regla general que hemos descubierto por inducción se refiere a dos números

yuxtapuestos de nuestra tabla, tales como 7 y 4, y con su suma, que es 11 en este caso. Ahora bien, en la sección anterior hemos visualizado la significación geométrica de 7, 4 y 11 en nuestro problema, y, al hacerlo así, hemos comprendido el *porqué* de $7 + 4 = 11$. Nos ocupamos con el paso de tres líneas a cuatro que dividen una superficie plana. Y al no haber ninguna particularidad especial en los números 3 y 4 pudimos pasar de cualquier número al siguiente, de $n$ a $n + 1$. El caso particular en discusión puede representarnos la situación general (ejemplo 2.10). Por nuestra parte, abandonaremos al lector el placer de extraer totalmente la idea general de la observación particular de la sección anterior. Al hacer esto encontrará una prueba formal de la regla descubierta inductivamente, por lo menos en lo que se refiere a las dos columnas.

Sin embargo, con objeto de completar la prueba debíamos considerar no sólo la disección de un plano por rectas, sino también la disección del espacio por planos. Sin embargo, esperamos que si hemos podido aclarar la disección de un plano la analogía nos ayudará a aclarar la disección del espacio. Nuevamente dejamos en esto al lector el placer de aprovecharse de la analogía.

## 18. Más conjeturas

Aún no hemos agotado el tema de las divisiones de planos y espacios. Hay algunos descubrimientos más por hacer y todos ellos son accesibles al razonamiento inductivo. Podemos fácilmente llegar a ello mediante la observación cuidadosa y la combinación inteligente de ejemplos particulares.

Podemos encontrar una *fórmula* para el número de divisiones de un plano por $n$ líneas en posición general. De hecho, ya tenemos una fórmula en un caso análogo más sencillo: $n$ puntos diferentes que dividen una línea recta en $n + 1$ segmentos. Esta fórmula análoga, los casos particulares propuestos en la tabla, la regla general descubierta inductivamente (que hemos casi probado), todos los resultados obtenidos hasta aquí pueden ayudarnos a resolver este nuevo problema. No entraré en detalles. Solamente apunto la solución que puede encontrarse de varias maneras, siguiendo las advertencias precedentes.

$1 + n$ es el número de puntos en que una línea recta es dividida por $n$ puntos diferentes.

$1 + n + \dfrac{n(n-1)}{2}$ es el número de partes en que un plano es dividido por $n$ líneas rectas en posición general.

El lector puede derivar la última fórmula o, al menos, puede comprobarla en los casos más sencillos, para $n = 0, 1, 2, 3, 4$. Igualmente abandonamos al lector el placer de descubrir una tercera fórmula de la misma clase para el número de divisiones del espacio. Al hacer este pequeño descubrimiento el lector puede ampliar su experiencia del razonamiento inductivo en cuestiones matemáticas y ver la ayuda que la analogía nos presta en la solución de problemas, grandes o pequeños.

## EJEMPLOS Y COMENTARIOS

La fórmula $C + V = A + 2$, conjeturada en la sección 1, es debida a Leonhard Euler. Es llamada «fórmula de Euler», mirada como una conjetura y considerada en diversos aspectos, unas veces inductivamente y otras con vista a encontrar una prueba, como en los ejemplos 1-10. Volvemos a ella en el ejemplo 21-30 y en el 31-41. Antes de intentar cualquier ejemplo en estas dos divisiones, leamos los ejemplos 21 y 31, respectivamente.

**1.** Dos pirámides, que permanecen con los lados opuestos y una base común, forman conjuntamente una «pirámide doble». Un octaedro es una pirámide doble particular; la base común es un cuadrado. ¿La fórmula de Euler es también para la pirámide doble general?

**2.** Tomemos un poliedro convexo con $C$ caras, $V$ vértices y $A$ aristas; elijamos un punto $P$ en su interior (su centroide, por ejemplo), describamos una esfera con centro $P$ y proyectemos el poliedro desde $P$ sobre la superficie de la esfera. Esta proyección transforma las $C$ caras en $R$ regiones o «países» sobre la superficie de la esfera, las $A$ aristas en una línea fronteriza que separa dos países vecinos y los $V$ vértices en una «esquina» o punto fronterizo común a tres o más países. Las líneas límites de esta proyección son de naturaleza particularmente sencilla (arcos de grandes círculos), pero es obvio que la validez de la fórmula de Euler para tal *subdivisión de la superficie de la esfera* en países es independiente de la precisión de las líneas límites; los números $C$, $V$ y $A$ son influidos por la continua deformación de estas líneas.

(1) Un *meridiano* es la mitad de un círculo que conecta los dos polos, Norte y Sur. Un *círculo paralelo* es la intersección de la superficie del globo con un plano paralelo al ecuador. La superficie de la Tierra está dividida en $m$ meridianos y $p$ paralelos en $R$ países. Calcular $C$, $V$ y $A$. ¿Es válida la fórmula de Euler?

(2) La proyección del octaedro desde su centro $P$ sobre la superficie de la esfera es un caso particular de la situación descrita en (1). ¿Para qué valores de $m$ y $p$?

**3.** La suerte juega un papel en los descubrimientos. Los descubrimientos inductivos dependen lógicamente de la observación física. En la sección 1 nos hemos encontrado con ciertos poliedros, pero podríamos cambiarlos por otros. Sin duda, no es que hayamos interpretado mal los cuerpos sólidos regulares, pero nuestra lista podía haber resultado de esta manera:

| *Poliedros* | $C$ | $V$ | $A$ |
|---|---|---|---|
| Tetraedro | 4 | 4 | 6 |
| Cubo | 6 | 8 | 12 |
| Octaedro | 8 | 6 | 12 |
| Prisma pentagonal | 7 | 10 | 15 |
| Pirámide doble pentagonal | 10 | 7 | 15 |
| Dodecaedro | 12 | 20 | 30 |
| Icosaedro | 20 | 12 | 30 |

¿Puede usted observar alguna regularidad? ¿Puede explicarla? ¿Cuál es la conexión con la fórmula de Euler?

**4.** Intentemos generalizar la relación entre dos poliedros observados en la tabla del ejemplo 3. [La relación descrita en la solución del ejemplo 3 bajo (2) es demasiado «estrecha», demasiado «detallada». Sin embargo, tomemos el cubo y el octaedro en la situación allí descrita, color rojo las aristas de uno, azul la del otro, y proyectémoslos desde su centro común $P$ sobre una esfera como la descrita en ejemplo 2. Luego, generalicemos.]

**5.** Sería suficiente probar la fórmula de Euler en un caso particular: para poliedros convexos que tienen solamente caras triangulares. ¿Por qué? [Sección 4.]

**6.** Sería suficiente probar la fórmula de Euler en un caso particular: para poliedros convexos que tienen sólo vértices de tres aristas. ¿Por qué? [Sección 4.]

**7.** Para probar la fórmula de Euler podemos restringirnos a figuras en un plano. En efecto, imaginemos que $C - 1$ caras del poliedro son hechas de cartón, mientras una es hecha de cristal; llamaremos a esta cara la «ventana». Miremos a través de la ventana en el interior del poliedro, manteniendo los ojos tan próximos a la ventana que vemos

todo el interior. (Esto sería imposible si el poliedro no fuese convexo.) Imaginemos que vemos una figura plana en el vidrio de la ventana: esto es, una *subdivisión* de la ventana en pequeños polígonos.

En esta subdivisión hay $N_2$ polígonos, $N_1$ líneas rectas límite (algunas internas y otras externas) y $N_0$ vértices.

(1) Expresar $N_0$, $N_1$ y $N_2$ en términos de $C$, $V$ y $A$.

(2) Si la fórmula de Euler se conserva para $C$, $V$ y $A$, ¿qué fórmula se conservará para $N_0$, $N_1$ y $N_2$?

**8.** Un rectángulo tiene $l$ pulgadas de largo y $m$ de ancho; $l$ y $m$ son enteros. El rectángulo está subdividido en $lm$ cuadrados iguales por líneas rectas paralelas a sus lados.

(1) Expresar $N_0$, $N_1$ y $N_2$ (definidos en el ejemplo 7) en términos de $l$ y $m$.

(2) ¿Es la relación ejemplo 7 (2) válida en el caso presente?

**9.** Los ejemplos 5 y 7 sugieren que examinemos la división de un triángulo en $N_2$ triángulos con $N_0$ — 3 vértices en el interior del triángulo subdividido. Al calcular la suma de todos los ángulos de $N_2$ triángulos de dos maneras diferentes, podemos probar la fórmula de Euler.

**10.** La sección 7 sugiere la extensión de la fórmula de Euler a cuatro y más dimensiones. ¿Cómo podemos hacer tangible una extensión tal? ¿Cómo podemos visualizarla?

La sección 7 muestra que el caso de los poliedros puede ser reducido a la subdivisión de un polígono plano. La analogía sugiere que el caso de cuatro dimensiones puede ser reducido a la subdivisión de un poliedro en nuestro visible espacio tridimensional. Si deseamos proceder inductivamente debemos examinar algunos ejemplos de tal subdivisión. Por analogía, el ejemplo 8 sugiere lo siguiente:

Una caja (que es un paralelepípedo rectangular) tiene las dimensiones de $l$, $m$ y $n$; estos tres números son enteros. La caja está subdividida en $lmn$ cubos iguales por planos paralelos a sus caras. $N_0$, $N_1$, $N_2$ y $N_3$ denotan el número de vértices, aristas, caras y los poliedros (cubos) que forman las subdivisiones, respectivamente.

(1) Expresar $N_0$, $N_1$, $N_2$ y $N_3$ en términos de $l$, $m$, y $n$.

(2) ¿Hay una relación análoga a la ecuación (2) en la solución del ejemplo 7?

**11.** $P_n$ denota el número de partes en que el plano es dividido por $n$ líneas rectas en posición general. Probar que

$$P_{n+1} = P_n + (n + 1).$$

**12.** $S_n$ denota el número de partes en que el espacio es dividido por $n$ planos en posición general. Probar que

$$S_{n+1} = S_n + P_n.$$

**13.** Verificar la fórmula conjetural

$$P_n = 1 + n + \frac{n(n-1)}{2}$$

para $n = 0, 1, 2, 3, 4$.

**14.** Intuir una fórmula para $S_n$ y verificarla para $n = 0, 1, 2, 3, 4, 5$.

**15.** ¿Cuántas partes son finitas de las once en que el plano se divide por cuatro líneas rectas en posición general? [¿Cuántas son infinitas?]

**16.** Generalizar los problemas anteriores.

**17.** ¿Cuántas partes son infinitas de las veintiséis en que es dividido el espacio por cinco planos en posición general?

**18.** Cinco planos pasan por el centro de la esfera, aunque en otros aspectos su división es general. Encontrar el número de partes en que la superficie de la esfera está dividida por cinco planos.

**19.** ¿En cuántas partes queda dividido el plano por cinco círculos que se intersectan mutuamente en posición general?

**20.** Generalicemos los problemas anteriores.

**21.** *Inducción: adaptación de la mente, adaptación del lenguaje.* La inducción termina por adaptar nuestra mente a los hechos. Cuando comparamos nuestras ideas con observaciones puede haber acuerdo o desacuerdo. Si hay acuerdo sentimos más confianza en nuestras ideas; si hay desacuerdo, las modificamos. Después de repetidas modificaciones nuestras ideas suelen adaptarse a los hechos mucho mejor. Nuestras primeras ideas sobre cualquier tema nuevo están muy cerca de ser erróneas, al menos en parte; el proceso inductivo nos da una oportunidad para corregirlas, adaptándolas a la realidad. Nuestros ejemplos muestran este proceso en pequeña escala, pero bastante claramente. En la sección 1, después de dos o tres conjeturas equivocadas, llegamos a una conjetura correcta. Podemos decir que llegamos a ella por accidente. «Sin embargo, tales accidentes sólo ocurren a la gente que los merece», como Lagrange dijo una vez cuando se discutió el más grande descubrimiento de Newton.

La adaptación de la mente puede ser más o menos lo mismo que la

adaptación del lenguaje; en cualquier caso, ambos van mano sobre mano. El progreso de la ciencia se nota en el progreso de la terminología. Cuando los físicos empezaron a hablar de «electricidad», o los médicos de «contagio», estos términos eran vagos, oscuros, confusos. Los términos que los científicos usan hoy, tales como «carga eléctrica», «corriente eléctrica», «infección de hongos», «infección de virus», son incomparablemente más claros y más definidos. Pero ¡qué cantidad de observación, cuántos ingeniosos experimentos yacen entre las dos terminologías y otros grandes descubrimientos! La inducción cambia la terminología, clarifica los conceptos. Nosotros podemos también ilustrar este aspecto del proceso, la clarificación inductiva de conceptos, mediante un conveniente ejemplo matemático a pequeña escala. La situación, no infrecuente en la investigación matemática, es ésta: Un teorema ha sido ya formulado, pero debemos darle un significado más preciso a los términos al objeto de hacerlos estrictamente correctos. Esto puede ser hecho por un proceso inductivo, como veremos.

Regresemos al ejemplo 2 y su solución. Allí hemos hablado de «subdivisión de la esfera en países» sin proponernos una definición formal de este término. Esperamos que la fórmula de Euler será válida si $C$, $V$ y $A$ denotan el número de países, líneas límite y esquinas en tal subdivisión. Sin embargo, nos fiamos de ejemplos y de una grosera descripción, sin dar definiciones formales de $C$, $V$ y $A$. ¿En qué exacto sentido habrá que tomar dichos términos para hacer estrictamente correcta la fórmula de Euler? Esa es nuestra cuestión.

Digamos que una subdivisión de la esfera (esto es, la *superficie* esférica) con una interpretación correspondiente a $C$, $V$ y $A$ es «correcta» si la fórmula de Euler se mantiene, y está «equivocada» si no se mantiene. Propongamos ejemplos de subdivisiones que nos ayuden a descubrir alguna distinción clara y sencilla entre los casos «correctos» y «equivocados».

**22.** La superficie total del globo es ocupada por un solo país. ¿Es esto correcto? (Queremos decir, correcto desde el punto de vista de la fórmula de Euler.)

**23.** La superficie del globo está dividida sólo en dos países, el hemisferio oriental y el occidental, separados por un gran círculo. ¿Es esto equivocado?

**24.** Los paralelos dividen la esfera en tres países. ¿Es ello correcto o equivocado?

**25.** Tres meridianos dividen la esfera en tres países. ¿Es correcto o equivocado?

**26.** Llamemos la división de la esfera por $m$ meridianos y $p$ círculos paralelos «división $(m, p)$»; cf. ejemplo 2 (1). ¿Es el caso extremo $(0, p)$ correcto o equivocado?

**27.** ¿Es el caso extremo $(m, 0)$ correcto o equivocado? (Cf. ejemplo 26.)

**28.** ¿Qué subdivisiones $(m, p)$ (cf. ejemplo 26) pueden ser generadas por el procedimiento descrito en el ejemplo 2? (Proyección de un poliedro convexo sobre la esfera, seguido por el desplazamiento de los límites del número de países y de fronteras alrededor de cada país inalterado.) ¿Qué condiciones concernientes a $m$ y $p$ caracterizan tales subdivisiones?

**29.** ¿Qué está equivocado en los ejemplos en que fracasa la fórmula de Euler? ¿Qué condiciones geométricas, haciendo más preciso el significado de $C$, $V$ y $A$, asegurarían la validez de la fórmula de Euler?

**30.** Proponer más ejemplos para ilustrar la contestación al ejemplo 29.

**31.** *La obra de Descartes sobre poliedros.* Entre los manuscritos dejados por Descartes se han encontrado breves notas sobre la teoría general de los poliedros. Una copia de estas notas (a mano de Leibniz) fue descubierta y publicada en 1860, más de doscientos años después de la muerte de Descartes; cf. las *Oeuvres* de Descartes, vol. 10, págs. 257-276. Estas notas tratan de temas íntimamente relacionados con el teorema de Euler; aunque las notas no contienen explícitamente el teorema contienen resultados de los que se sigue inmediatamente.

Consideremos, con Descartes, un poliedro convexo. Llamemos a cualquier ángulo de una cualquiera de las caras del poliedro *ángulo superficie*, y establezcamos $\Sigma\alpha$ para la suma de todos los ángulos superficie. Descartes calculó $\Sigma\alpha$ de dos maneras diferentes, y el teorema de Euler resulta de la comparación de esas dos expresiones.

Los ejemplos siguientes dan al lector una oportunidad para reconstruir algunas conclusiones de Descartes. He aquí la notación que utilizaremos:

$C_n$ denota el número de caras con $n$ aristas,

$V_n$, el número de vértices en que $n$ aristas terminan, de modo que

$$C_3 + C_4 + C_5 + \ldots = C,$$
$$V_3 + V_4 + V_5 + \ldots = V.$$

Por lo demás, continuaremos llamando $A$ al número total de aristas del poliedro.

**32.** Expresar el número de todos los ángulos superficie de tres maneras diferentes: en términos de $C_3$, $C_4$, $C_5$, ..., de $V_3$, $V_4$, $V_5$, ..., y de $A$, respectivamente.

**33.** Calcular $\Sigma\alpha$ para los cinco sólidos regulares: el tetraedro, el cubo, el octaedro, el dodecaedro, el icosaedro.

**34.** Expresar $\Sigma\alpha$ en términos de $C_3$, $C_4$, $C_5$, ...

**35.** Expresar $\Sigma\alpha$ en términos de $A$ y $C$.

**36.** *Angulos sólidos suplementarios y polígonos esféricos suplementarios.* Llamamos *ángulo sólido* a lo que usualmente se conoce como ángulo poliédrico.

Dos ángulos sólidos convexos tienen el mismo número de caras y de vértices, pero ningún otro punto en común. A cada cara de un ángulo sólido corresponde una arista del otro, y la cara es perpendicular a la arista correspondiente. (Esta relación entre los dos ángulos sólidos es recíproca: la arista $a$, intersección de dos caras contiguas del primer ángulo sólido, corresponde a la cara $c'$, del segundo ángulo sólido, si $c'$ es limitada por las dos aristas correspondientes a las dos caras mencionadas arriba.) Dos ángulos sólidos en esta relación mutua son llamados ángulos sólidos *suplementarios*. (Este nombre no es corriente, pero los dos ángulos suplementarios ordinarios pueden ser llevados a una posición análoga mutua.) Cada uno de los dos ángulos sólidos suplementarios recibe el nombre de suplemento del otro.

La esfera con radio 1, descrita sobre el vértice común de dos ángulos sólidos suplementarios como centro, se interseccíona con éstos en dos polígonos esféricos: también estos polígonos reciben el nombre de suplementarios.

Consideremos dos polígonos esféricos suplementarios, donde $a_1$, $a_2$, ... $a_n$ denotan los lados del primer polígono, $\alpha_1$, $\alpha_2$, ... $\alpha_n$ sus ángulos, $A$ su área, $P$ su perímetro, y $a'_1$, $a'_2$, ... $a'_n$, $\alpha'_1$, $\alpha'_2$, ... $\alpha'_n$, $A'$, $P'$ son las partes análogas del otro polígono. Luego, si la notación es apropiadamente elegida,

$$a_1 + \alpha_1 = a_2 + \alpha'_2 = \ldots = a_n + \alpha'_n = \pi,$$
$$a'_1 + \alpha_1 = a'_2 + \alpha_2 = \ldots = a'_n + \alpha_n = \pi;$$

esto es muy conocido y fácilmente visto.

Probemos que

$$P + A' = P' + A = 2\pi.$$

[Se supone que el lector conoce que el área de un triángulo esférico con ángulos $\alpha$, $\beta$ y $\gamma$ es el «exceso esférico» $\alpha + \beta + \gamma - \pi$.]

**37.** «Como en una figura plana todos los ángulos exteriores juntos son iguales a cuatro ángulos rectos, del mismo modo en una figura sólida todos los ángulos sólidos exteriores son iguales a ocho ángulos rectos.» Intentar interpretar esta frase encontrada en las notas de Descartes como un teorema que puede ser probado. [Ver la fig. 3.7.]

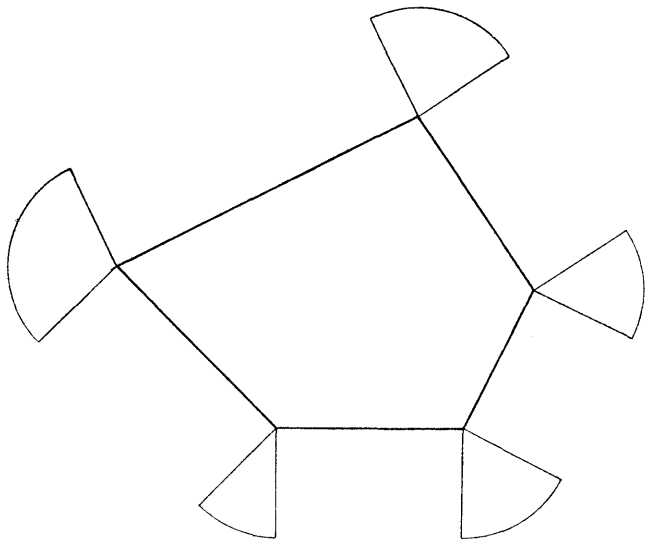


Fig. 3.7. Angulos exteriores de un polígono.

**38.** Expresar $\Sigma\alpha$ en términos de $V$.

**39.** Probar el teorema de Euler.

**40.** La observación inicial de la sección 1 es vaga, pero puede sugerir varias proposiciones precisas. Hay una que no hemos considerado en la sección 1: «Si cualquiera de las tres cantidades $C$, $V$ y $A$ tiende a $\infty$, también las otras dos deben tender a $\infty$». Probar las desigualdades siguientes que se conservan generalmente para los poliedros convexos y dan aún información más precisa:

$$2A \geqq 3C, \quad 2V \geqq C + 4, \quad 3V \geqq A + 6,$$
$$2A \geqq 3V, \quad 2C \geqq V + 4, \quad 3C \geqq A + 6.$$

¿Puede el caso de igualdad ser considerado desde estas desigualdades? ¿Qué clases de poliedros pueden ser considerados desde ellas?

**41.** Hay poliedros convexos cuyas caras son polígonos de la misma

clase, es decir, polígonos con el mismo número de lados. Por ejemplo, todas las caras de un tetraedro son triángulos; todas las de un paralelepípedo, cuadriláteros; todas las de un dodecaedro regular, pentágonos. «Y así sucesivamente», podemos ser tentados a decir. Sin embargo, tal inducción puede ser desorientadora: *no* hay poliedros convexos con caras que sean todas hexágonos. Intentar probarlo. [Ejemplo 31.]

# La inducción en la teoría de los números

*En la teoría de los números sucede con bastante frecuencia que las verdades más bellas brotan por inducción.*—GAUSS *.

## 1. TRIÁNGULOS RECTÁNGULOS EN ENTEROS **.

El triángulo con lados 3, 4 y 5 es un triángulo rectángulo, puesto que

$$3^2 + 4^2 = 5^2.$$

Este es el ejemplo más sencillo de un triángulo rectángulo en que los lados son medidos por enteros. Tales «triángulos rectángulos en enteros» han jugado un papel en la historia de la teoría de los números; ya los antiguos babilonios descubrieron algunas de sus propiedades.

Uno de los más obvios problemas sobre tales triángulos es el siguiente: *¿Hay un triángulo rectángulo en enteros cuya hipotenusa es un número entero dado n?*

Concentrémonos sobre este problema. Buscamos un triángulo cuya hipotenusa mide el entero $n$ y sus lados los enteros $x$ e $y$. Podemos asumir que $x$ denota el más largo de los dos lados. Por tanto, estando dado $n$, se pueden buscar dos enteros $x$ e $y$, tales que

$$n^2 = x^2 + y^2, \qquad 0 < y \leq x < n.$$

Podemos atacar el problema inductivamente y, a menos que tengamos algún conocimiento completamente específico, no podemos enfrentarnos a él de otra manera. Pongamos un ejemplo. Hagamos $n = 12$. Por tanto, busquemos dos enteros positivos $x$ e $y$, tales que $x \geq y$

$$144 = x^2 + y^2.$$

---

* *Werke*, vol. 2, pág. 3.

** Partes de este capítulo han aparecido ya bajo el título «Enseñemos a intuir», en el volumen *Etudes de Philosophie des Sciences en hommage à Ferdinand Gonseth.* Editions du Griffon, 1950; véanse págs. 147-154.

¿Qué valores son convenientes para $x^2$? Los siguientes:

$$1, 4, 9, 16, 25, 36, 49, 64, 81, 100, 121.$$

¿Es $x^2 = 121$? Es decir, ¿es

$$144 - x^2 = 144 - 121 = y^2$$

un cuadrado? No, 23 no es un cuadrado. Intentemos ahora otros cuadrados; de hecho no necesitamos intentar demasiado, ya que $y \leq x$,

$$144 = x^2 + y^2 \leq 2x^2$$

$$x^2 \geq 72.$$

Por tanto, $x^2 = 100$ y $x^2 = 81$ son las únicas posibilidades que quedan. Ahora bien, ninguno de los números

$$144 - 100 = 44, \quad 144 - 81 = 63$$

son cuadrados y de aquí la contestación: no hay ningún triángulo rectángulo en enteros con hipotenusa 12.

Hagamos lo mismo con la hipotenusa 13. De los tres números

$$169 - 144 = 25, \quad 169 - 121 = 48, \quad 169 - 100 = 69$$

solamente uno de ellos es cuadrado, de modo que sólo hay un triángulo rectángulo en enteros que tenga de hipotenusa 13:

$$169 = 144 + 25.$$

Procediendo de modo semejante podemos examinar con un poco de paciencia todos los números bajo uno dado, que no sea demasiado alto; por ejemplo, 20. Encontramos únicamente cinco «hipotenusas» menores de 20, los números 5, 10, 13, 15 y 17:

$$\begin{aligned} 25 &= 16 + 9 \\ 100 &= 64 + 36 \\ 169 &= 144 + 25 \\ 225 &= 144 + 81 \\ 289 &= 225 + 64. \end{aligned}$$

A propósito, los casos 10 y 15 no son muy interesantes. El triángulo con lados 10, 8 y 6 es semejante al triángulo más sencillo de lados 5, 4 y 3, y lo mismo ocurre con el triángulo de lados 15, 12 y 9. Los tres triángulos rectángulos que quedan de hipotenusa 5, 13 y 17, respectivamente, son esencialmente diferentes; ninguno de ellos son semejantes entre s .

Podemos observar que los tres números 5, 13 y 17 son *primos impares*. Sin embargo, ésos no son todos los *primos impares* que están por debajo de 20; ninguno de los otros primos impares, 3, 7, 11 y 19, son hipotenusas. ¿Por qué es esto? ¿Cuál es la diferencia entre las dos series? *¿Cuándo, bajo qué circunstancias, es un primo impar la hipotenusa de un triángulo rectángulo en enteros, y cuándo no?*

Esto es una modificación de nuestro problema original. Al pronto, parece más esperanzador; por lo menos, es algo nuevo. Investiguémoslo —otra vez inductivamente—. Con un poco de paciencia construyamos la siguiente tabla (el guión indica que no hay triángulo rectángulo con hipotenusa $p$).

| *Primo impar p* | *Triángulos rectángulos con hipotenusa p* |
|---|---|
| 3 | — |
| 5 | $25 = 16 + 9$ |
| 7 | — |
| 11 | — |
| 13 | $169 = 144 + 25$ |
| 17 | $289 = 225 + 64$ |
| 19 | — |
| 23 | — |
| 29 | $841 = 441 + 400$ |
| 31 | — |

¿Cuándo es un número primo una hipotenusa; cuándo no lo es? ¿Cuál es la diferencia entre ambos casos? Un físico se preguntaría a sí mismo algunas cuestiones semejantes. Por ejemplo, él investiga la doble refracción de los cristales. Algunos de éstos muestran doble refracción; otros no. ¿Qué cristales son de doble refracción y cuáles no? ¿Cuál es la diferencia entre los dos casos?

El físico mira los cristales, y nosotros miramos nuestras dos series de números primos

$$5, 13, 17, 29, \ldots \quad \text{y} \quad 3, 7, 11, 19, 23, 31, \ldots$$

Por nuestra parte, buscamos algunas diferencias características entre ambas series. En las dos, los números primos aumentan a saltos irregulares. Miremos la longitud de estos saltos y las diferencias sucesivas:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | | 13 | | 17 | | 29 | 3 | | 7 | | 11 | | 19 | | 23 | | 31 |
| | 8 | | 4 | | 12 | | | 4 | | 4 | | 8 | | 4 | | 8 | |

Muchas de estas diferencias son iguales a 4 y, como es fácil observar, *todas son divisibles por* 4. Los números primos de la primera serie, que comienza por 5, dejan el resto 1 cuando se divide por 4 y son de la forma $4n + 1$ con integral $n$. Los números primos de la segunda serie, que empiezan por 3, son de la forma $4n + 3$. ¿Podría ser ésta la diferencia característica que estamos buscando? Si no descartamos esta posibilidad desde el principio, nos vemos conducidos a la conjetura siguiente: *Un número primo de la forma* $4n + 1$ *es la hipotenusa de un triángulo rectángulo en enteros; un número primo de la forma* $4n + 3$ *no es la hipotenusa del triángulo.*

## 2. Sumas de cuadrados.

El problema de los triángulos rectángulos en enteros, uno de cuyos aspectos acabamos de discutir en la sección 1, jugó, como hemos dicho, un papel importante en la historia de la teoría de los números. En efecto, continúa planteándonos numerosas cuestiones. ¿Qué números, sean o no cuadrados, se pueden descomponer en dos cuadrados? ¿Qué ocurre con los números que no se pueden descomponer en dos cuadrados? Quizá se pueden descomponer en tres cuadrados; pero ¿qué ocurre con los números que no pueden ser descompuestos en tres cuadrados?

Podríamos seguir indefinidamente. Bachet de Méziriac (autor del primer libro impreso sobre recreaciones matemáticas) notó que *cualquier número* (que sea un entero positivo) *es, o un cuadrado o la suma de dos, tres o cuatro cuadrados.* El no pretendió poseer una prueba. Encontró indicaciones que apuntaban a esta proposición en algunos problemas de Diofanto y la verificó hasta el 325.

En resumen, la proposición de Bachet era una conjetura encontrada inductivamente. Me parece que su principal logro fue plantearse la cuestión: *¿Cuántos cuadrados se necesitan para representar a todos los enteros?* Una vez que esta cuestión se ha planteado con claridad no es

muy difícil descubrir la respuesta inductivamente. Construyamos una tabla que empieza con

$$\begin{aligned}
1 &= 1 \\
2 &= 1 + 1 \\
3 &= 1 + 1 + 1 \\
4 &= 4 \\
5 &= 4 + 1 \\
6 &= 4 + 1 + 1 \\
7 &= 4 + 1 + 1 + 1 \\
8 &= 4 + 4 \\
9 &= 9 \\
10 &= 9 + 1.
\end{aligned}$$

Esto verifica la conjetura hasta 10. Solamente el número 7 requiere cuatro cuadrados; los otros son representables por uno, dos o tres. Bachet llegó a tabular hasta 325 y encontró muchos números que requerían cuatro cuadrados, pero ninguno más. Tal evidencia inductiva le satisfizo, al menos hasta cierto grado, y por ello publicó su conjetura. Tuvo suerte. Su conjetura era cierta y así llegó a establecer el «teorema de los cuatro cuadrados», que formuló de esta manera: La ecuación

$$n = x^2 + y^2 + z^2 + w^2,$$

donde $n$ es un número entero positivo dado, tiene una solución en la que $x$, $y$, $z$ y $w$ son enteros no-negativos.

La descomposición de un número en una suma de cuadrados tiene aún otros aspectos. Así, podemos investigar el número de soluciones de la ecuación

$$n = x^2 + y^2$$

de enteros $x$ e $y$. Podemos admitir sólo enteros positivos, o toda clase de enteros: positivos, negativos y 0. Si elegimos la última concepción del problema y tomamos como ejemplo $n = 25$, encontramos doce soluciones de la ecuación

$$25 = x^2 + y^2,$$

es decir, las siguientes:

$$\begin{aligned}
25 &= 5^2 + 0^2 = (-5)^2 + 0^2 = 0^2 + 5^2 = 0^2 + (-5)^2 \\
&= 4^2 + 3^2 = (-4)^2 + 3^2 = 4^2 + (-3)^2 = (-4)^2 + (-3)^2 \\
&= 3^2 + 4^2 = (-3)^2 + 4^2 = 3^2 + (-4)^2 = (-3)^2 + (-4)^2.
\end{aligned}$$

A propósito, estas soluciones tienen una interesante interpretación geométrica, pero ahora no la examinaremos. Ver ejemplo 2.

## 3. Sobre la suma de cuatro cuadrados impares

De los muchos problemas que se refieren a las sumas de cuadrados hemos elegido uno que parecía de alcance lejano, pero que puede volverse excepcionalmente instructivo.

*Denotemos con u un número entero, positivo e impar. Investiguemos el número de soluciones a la ecuación*

$$4u = x^2 + y^2 + z^2 + w^2,$$

*en que x, y, z y w son enteros positivos impares.*

Por ejemplo, si $u = 1$, tenemos la ecuación

$$4 = x^2 + y^2 + z^2 + w^2,$$

donde hay obviamente una solución única, $x = y = z = w = 1$. En efecto, no podemos considerar

$$x = -1, \quad y = 1, \quad z = 1, \quad w = 1$$

o

$$x = 2, \quad y = 0, \quad z = 0, \quad w = 0$$

como solución, puesto que hemos admitido sólo números impares positivos para $x, y, z$ y $w$. Si $u = 3$, la ecuación será

$$12 = x^2 + y^2 + z^2 + w^2$$

y las dos soluciones siguientes:

$$x = 3, \quad y = 1, \quad z = 1, \quad w = 1$$
$$x = 1, \quad y = 3, \quad z = 1, \quad w = 1,$$

son diferentes.

Con objeto de enfatizar la restricción impuesta sobre los valores de $x, y, z$ y $w$ evitaremos el término «solución» y usaremos la descripción más específica: «representación de $4u$ como una suma de cuatro cuadrados impares». Como esta descripción es muy larga, la abreviaremos de diversas maneras, algunas con sólo la palabra «representación».

## 4. Examinando un ejemplo

Con objeto de familiarizarnos con el significado de nuestro problema consideremos un ejemplo. Elijamos $u = 25$. Luego, $4u = 100$ y hallemos todas las representaciones de 100 como una suma de cuatro cuadrados impares. ¿Qué cuadrados impares se pueden obtener para este propósito? Los siguientes:

$$1, 9, 25, 49, 81.$$

Si 81 es uno de los cuatro cuadrados cuya suma es 100, la suma de los otros tres será

$$100 - 81 = 19.$$

Los únicos cuadrados impares menores de 19 son 1 y 9, y $19 = 9 + 9 + 1$ es, evidentemente, la sola posibilidad de representar 19 como una suma de tres cuadrados impares si los términos son colocados en orden de magnitud. Así, obtenemos:

$$100 = 81 + 9 + 9 + 1.$$

De modo semejante, encontramos

$$\begin{aligned} 100 &= 49 + 49 + 1 + 1, \\ 100 &= 49 + 25 + 25 + 1, \\ 100 &= 25 + 25 + 25 + 25. \end{aligned}$$

Procediendo sistemáticamente, por separación de los primeros cuadrados más grandes, podemos convencernos de que hemos agotado todas las posibilidades, teniendo en cuenta que los cuatro cuadrados son colocados en orden descendiente (o más bien, en orden no ascendiente). Pero hay más posibilidades, si tenemos en cuenta todas las combinaciones de los términos. Por ejemplo,

$$\begin{aligned} 100 &= 49 + 49 + 1 + 1 \\ 100 &= 49 + 1 + 49 + 1 \\ 100 &= 49 + 1 + 1 + 49 \\ 100 &= 1 + 49 + 49 + 1 \\ 100 &= 1 + 49 + 1 + 49 \\ 100 &= 1 + 1 + 49 + 49. \end{aligned}$$

Estas seis sumas tienen los mismos términos, pero cambia su orden; pueden ser consideradas, de acuerdo con el planteamiento de nuestro problema, como seis representaciones diferentes; una representación

$$100 = 49 + 49 + 1 + 1$$

con términos que no incrementan puede ser considerada como fuente de otras cinco representaciones, es decir, de seis representaciones en total. Tenemos, pues,

| *Términos que no incrementan* | *Número de ordenaciones* |
|---|---|
| $81 + 9 + 9 + 1$ | 12 |
| $49 + 49 + 1 + 1$ | 6 |
| $49 + 25 + 25 + 1$ | 12 |
| $25 + 25 + 25 + 25.$ | 1 |

Sumando, encontramos, en nuestro caso, donde $u = 25$ y $4u = 100$, que

$$12 + 6 + 12 + 1 = 31$$

representaciones de $4u = 100$ como una suma de cuatro cuadrados impares.

## 5. Tabulando las observaciones

El caso particular $u = 25$ donde $4u = 100$ y el número de representaciones, 31, nos ha mostrado claramente el significado del problema. Podemos ahora explorar sistemáticamente los casos más sencillos, $u = 1, 3, 5, \ldots$ hasta $u = 25$. Construyamos la tabla I. (Véase en la página siguiente; el lector construirá la tabla por sí mismo, o, al menos, la verificará en algunos términos.)

TABLA I

| $u$ | $4u$ | *No incrementan* | *Ordenaciones* | *Representaciones* |
|---|---|---|---|---|
| 1 | 4 | $1+1+1+1$ | 1 | 1 |
| 3 | 12 | $9+1+1+1$ | 4 | 4 |
| 5 | 20 | $9+9+1+1$ | 6 | 6 |
| 7 | 28 | $25+1+1+1$ | 4 | 8 |
| | | $9+9+9+1$ | 4 | |
| 9 | 36 | $25+9+1+1$ | 12 | 13 |
| | | $9+9+9+9$ | 1 | |
| 11 | 44 | $25+9+9+1$ | 12 | 12 |
| 13 | 52 | $49+1+1+1$ | 4 | 14 |
| | | $25+25+1+1$ | 6 | |
| | | $25+9+9+9$ | 4 | |
| 15 | 60 | $49+9+1+1$ | 12 | 24 |
| | | $25+25+9+1$ | 12 | |
| 17 | 68 | $49+9+9+1$ | 12 | 18 |
| | | $25+25+9+9$ | 6 | |
| 19 | 76 | $49+25+1+1$ | 12 | 20 |
| | | $49+9+9+9$ | 4 | |
| | | $25+25+25+1$ | 4 | |
| 21 | 84 | $81+1+1+1$ | 4 | 32 |
| | | $49+25+9+1$ | 24 | |
| | | $25+25+25+9$ | 4 | |
| 23 | 92 | $81+9+1+1$ | 12 | 24 |
| | | $49+25+9+9$ | 12 | |
| 25 | 100 | $81+9+9+1$ | 12 | 31 |
| | | $49+49+1+1$ | 6 | |
| | | $49+25+25+1$ | 12 | |
| | | $25+25+25+25$ | 1 | |

## 6. ¿CUÁL ES LA REGLA?

¿Hay alguna ley que pueda ser reconocida, cualquier sencilla conexión entre el número impar $u$ y el número de diferentes representaciones de $4u$ como una suma de cuatro cuadrados impares?

Esta cuestión es el meollo de nuestro problema. Tenemos que contestarla sobre la base de las observaciones recogidas y tabuladas en la sección anterior. Estamos en la posición del naturalista intentando sacar alguna regla, alguna fórmula general de sus datos experimentales.

Nuestro material experimental accesible en este momento consiste en dos series paralelas de números:

1 3 5 7 9 11 13 15 17 19 21 23 25
1 4 6 8 13 12 14 24 18 20 32 24 31.

La primera serie es la sucesión de los números impares, pero ¿cuál es la regla que gobierna la segunda serie?

Cuando intentamos contestar esta cuestión, nuestro primer sentimiento puede ser encerrarnos en la desesperación. La segunda serie parece completamente irregular, su complejo origen nos desazona y la esperanza de encontrar alguna regla se nos va. Sin embargo, si nos olvidamos de este complejo origen y nos concentramos en lo que está ante nosotros, hay un punto muy fácil de notar. Sucede bastante a menudo que un término de la primera serie excede al término correspondiente de la segunda serie justo en una unidad. Señalemos estos casos con tipos gruesos y la serie de nuestro material experimental presentará el siguiente aspecto:

**1 3 5 7** 9 **11 13** 15 **17 19** 21 **23** 25
1 4 6 8 13 12 14 24 18 20 32 24 31.

Los números en tipos gruesos atraen nuestra atención. No es difícil reconocerlos: son números *primos*. En efecto, son *todos* los primos de la primera fila hasta el fin de nuestra tabla. Esta observación parecerá muy sorprendente si recordamos el origen de la serie. Solamente considerábamos entonces los cuadrados, sin referencia ninguna a los primos. ¿No es extraño que los números primos jueguen un papel en nuestro problema? Es difícil evitar la impresión de que nuestra observación es significativa, de que hay algo digno de notar tras ella.

¿Qué pasa con los números de la primera serie que no están en tipos gruesos? Se trata de números impares, aunque no primos. El primero, 1, es la unidad; los otros son compuestos:

$$9 = 3 \times 3, \quad 15 = 3 \times 5, \quad 21 = 3 \times 7, \quad 25 = 5 \times 5.$$

¿Cuál es la naturaleza de los números correspondientes en la segunda serie?

Si el número impar $u$ es un primo, el número correspondiente será $u + 1$; si $u$ no es primo, el correspondiente no será $u + 1$. Esto es lo observado hasta aquí; ahora podemos añadir una ligera nota. Si $u = 1$,

el número correspondiente es también 1, y por ello *menos* que $u + 1$, pero en todos los casos restantes en que $u$ no es primo el número correspondiente es siempre *más grande* que $u + 1$. Es decir, el número correspondiente a $u$ es menor, igual, o mayor que $u + 1$ según que $u$ sea la unidad, un número primo o uno compuesto. Existe, pues, alguna regularidad.

Concentrémonos sobre los números compuestos en la línea superior y los correspondientes de la más baja:

| $3 \times 3$ | $3 \times 5$ | $3 \times 7$ | $5 \times 5$ |
|---|---|---|---|
| 13 | 24 | 32 | 31 |

Aquí hay algo extraño. Los cuadrados en la primera línea corresponden a los primos en la segunda línea. Sin embargo, aún tenemos pocas observaciones; probablemente no daremos demasiado peso a esta observación. Pues, inversamente de lo visto hasta aquí, sigue siendo cierto que, bajo los números compuestos de la primera línea que no son cuadrados, encontramos números de la segunda fila que no son primos:

| $3 \times 5$ | $3 \times 7$ |
|---|---|
| $4 \times 6$ | $4 \times 8$ |

De nuevo vemos aquí algo extraño. Cada factor de la segunda línea excede en una unidad al factor correspondiente de la primera línea. Seguimos con demasiado pocas observaciones; lo mejor es no conceder mucho peso a este síntoma. Por el contrario, observemos antes

$$\begin{matrix} p \\ p + 1 \end{matrix}$$

y notemos que

$$\begin{matrix} pq \\ (p + 1)\ \ (q + 1), \end{matrix}$$

donde $p$ y $q$ son primos. Hay en ello cierta regularidad.

Quizá lo veremos más claramente si escribimos la anotación $pq$ de modo diferente:

$$(p + 1)\,(q + 1) = pq + p + q + 1.$$

¿Qué podemos ver aquí? ¿Qué son estos números $pq, p, q, 1$? De cualquier manera, los casos

| | |
|---|---|
| 9 | 25 |
| 13 | 31 |

permanecen inexplicados. De hecho, las anotaciones correspondientes a 9 y 25 son mayores que 9 + 1 y 25 + 1, respectivamente, como hemos ya observado:

$$13 = 9 + 1 + 3 \qquad 31 = 25 + 1 + 5.$$

¿Qué son estos números?

Si algo de luz nos llega de alguna parte, podremos llegar a combinar nuestras fragmentarias observaciones en un todo coherente, nuestras indicaciones desparramadas dentro de una visión iluminativa de la correspondencia total:

| $p$ | $pq$ | 9 | 25 | 1 |
|---|---|---|---|---|
| $p + 1$ | $pq + p + q + 1$ | $9 + 3 + 1$ | $25 + 5 + 1$ | 1. |

¡DIVISORES! La segunda línea exhibe los divisores de los números de la primera línea. Esta puede ser la regla deseada, y un descubrimiento, un descubrimiento real: *A cada número de la primera línea corresponde la suma de sus divisiones.*

Y así nos hemos visto conducidos a una conjetura, quizá a una de «las verdades más bellas» de que habla Gauss: *Si u es un número impar, el número de representaciones de 4u como suma de cuatro cuadrados impares es igual a la suma de los divisores de u.*

## 7. Sobre la naturaleza del descubrimiento inductivo

Regresemos a las secciones anteriores (3 a 6) y encontraremos algunas preguntas que hacer.

*¿Qué hemos obtenido?* Ni una prueba, ni la sombra de una prueba, sino una simple conjetura: una descripción sencilla de los hechos dentro de nuestro material experimental, y una cierta esperanza de que esta descripción pueda aplicarse más allá de sus límites.

*¿Cómo hemos obtenido la conjetura?* De la misma manera que la gente corriente, o los científicos que trabajan en campos no matemáticos, obtienen los suyos. Hemos recogido observaciones relevantes, las hemos examinado y comparado, hemos notado regularidades fragmentarias, dudado, desacertado, y eventualmente hemos logrado *combinar los detalles desparramados en un conjunto aparentemente significativo.* De modo muy semejante, un arqueólogo puede reconstruir el conjunto de una inscripción sobre la base de unas cuantas letras

desparramadas en una gastada piedra, o un paleontólogo la fisonomía de un animal extinguido por unos pocos de sus huesos petrificados. En nuestro caso, el conjunto significativo aparece en el momento en que reconocemos el concepto unificante apropiado (los divisores).

## 8. Sobre la naturaleza de la evidencia inductiva

Aún quedan algunas cuestiones más.

*¿Qué fuerza tiene la evidencia?* La cuestión es incompleta. Nos referimos, por supuesto, a la evidencia inductiva de la conjetura establecida en la sección 6 que hemos derivado de la tabla I de la sección 5; esto se comprende. Pero, ¿qué significamos por «fuerza»? La evidencia será fuerte si es convincente, y es convincente si convence a alguien. Sin embargo, con ello no decimos a quién debe convencer —¿a mí, a usted, a Euler, al principiante, a quién?

Personalmente, yo encuentro la evidencia bastante convincente. Estoy seguro de que Euler habría pensado bien de ella. (Menciono a Euler porque él estuvo muy cerca de descubrir nuestra conjetura; ver el ejemplo 6.24.) También creo que un principiante que conozca algo sobre la divisibilidad de los números debería encontrar la evidencia bastante convincente. Un colega mío, un excelente matemático que, sin embargo, no está familiarizado con este reducto de la teoría de los números, encontró la evidencia «cien por cien convincente».

*Yo no me refiero a impresiones subjetivas. ¿Cuál es el grado preciso, objetivamente evaluado de creencia racional, justificado por la evidencia inductiva?* Usted me da una cosa (A), fracasa en darme otra (B) y me pide una tercera (C).

(A) Usted me da exactamente la evidencia inductiva: la conjetura ha sido verificada en los primeros trece casos, para los números 4, 12, 20, ..., 100. Esto es perfectamente claro.

(B) Usted desea evaluar el grado de creencia racional justificado por esta evidencia. Sin embargo, tal creencia puede depender, si no de los caprichos y el temperamento, al menos del *conocimiento* de la persona que recibe la evidencia. El puede conocer una prueba del teorema conjetural o un contraejemplo que lo refuta. En ambos casos, el grado de su creencia, ya establecida firmemente, permanecerá inalterado por la evidencia inductiva. Pero si él conoce algo cercano a una prueba completa o una completa refutación del teorema, su creencia es todavía capaz de modificación y se verá afectada por la evidencia inductiva aquí producida, aunque de ello resulten grados diferentes

de creencia según la clase de conocimiento que tenga. Por tanto, si usted desea una contestación definida, usted debe especificar un nivel definido de conocimiento sobre el que la propuesta evidencia inductiva (A) sea juzgada. Usted me dará un conjunto definido de hechos relevantes (quizá una lista explícita de conocidas proposiciones elementales en la teoría de los números).

(C) Usted desea evaluar exactamente el grado de creencia racional justificado por la evidencia inductiva. ¿Podría dárselo quizá expresado en porcentajes de «creencia completa»? (Pongámonos de acuerdo en llamar «creencia completa» al grado de creencia justificada por una completa prueba matemática del teorema en cuestión.) ¿Usted esperaría que yo le dijese que la evidencia dada justifica una creencia al 99 por 100 o el 2.875 por 100, o el .000001 por 100 de la «creencia completa»?

En resumen, usted desea resolver un problema: Dada la evidencia inductiva (A) y un definido conjunto de hechos o proposiciones conocidas (B), calcular el porcentaje de creencia completa racionalmente resultante de ambos (C).

Resolver este problema es mucho más de lo que yo puedo hacer. No conozco a nadie que pueda hacerlo ni que siquiera se atreva a hacerlo. Sé de algunos filósofos que pretendieron hacer algo de esta clase a grandes rasgos. Sin embargo, una vez enfrentados con el problema concreto se evaden y defienden, encontrando mil excusas para no atacarlo.

Quizá el problema es uno de esos típicos problemas filosóficos sobre los que se puede hablar y hablar en general, y que llegan a inquietarnos realmente, pero que se desvanecen en la nada cuando tratamos de llevarlos a términos concretos.

*¿No se podría comparar el presente caso de inferencia inductiva con algún caso típico y llegar así a una estimación razonable de la fortaleza de su evidencia?* Comparemos la evidencia inductiva de nuestra conjetura con la evidencia de Bachet para la suya.

La conjetura de Bachet era: Para $n = 1, 2, 3, \ldots$ la ecuación

$$n = x^2 + y^2 + z^2 + w^2$$

tiene al menos una solución en enteros no-negativos $x$, $y$, $z$ y $w$. Verificó esta conjetura para $n = 1, 2, 3, \ldots, 325$. (Ver la sección 2, especialmente la tabla corta.)

Nuestra conjetura es: Para un número impar dado $u$, el número de soluciones de la ecuación

$$4u = x^2 + y^2 + z^2 + w^2$$

en enteros impares positivos $x, y, z$ y $w$ es igual a la suma de los divisores de $u$. Hemos verificado esta conjetura para $u = 1, 3, 5, 7, \ldots, 25$ (trece casos). (Ver las secciones 3 a 6.)

Ahora compararé estas dos conjeturas y la evidencia inductiva nos dará el resultado en tres aspectos, según las respectivas verificaciones.

*Número de verificaciones.* La conjetura de Bachet se verificó en trescientos veinticinco casos, mientras la nuestra sólo en trece; la de Bachet tiene por este lado una clara ventaja.

*Precisión de predicción.* La conjetura de Bachet predice que el número de soluciones es $\geq 1$; la nuestra predice que el número de soluciones es exactamente igual a tal y tal cantidad. Es razonable suponer, en mi opinión, que *la verificación de una predicción más precisa le da más peso* que la de una predicción menos precisa. La ventaja en este aspecto cae claramente de nuestro lado.

*Conjeturas rivales.* La conjetura de Bachet se refiere al máximo número de cuadrados, digamos $M$, que se necesitan para representar un entero positivo como suma de cuadrados. En efecto, la conjetura de Bachet asegura que $M = 4$. Yo no creo que Bachet tuviese ninguna razón *a priori* para preferir $M = 4$ a, digamos, $M = 5$, o cualquier otro valor, como $M = 6$ o $M = 7$; incluso $M = \infty$ no es excluido *a priori*. (Naturalmente, $M = \infty$ significaría que hay más y más grandes enteros pidiendo más y más cuadrados. Por otro lado, $M = \infty$ aparecerá como la conjetura más plausible.) En resumen, la conjetura de Bachet tiene muchas rivales. Pero la nuestra no tiene ninguna. Mirando la secuencia irregular de los números de representaciones (sección 6) teníamos la impresión de no poder establecer ninguna regla. Ahora encontramos una regla muy clara y esperamos encontrar cualquiera otra más.

Puede ser difícil escoger una novia donde hay muchas bellas señoritas; si sólo existe una muchacha a nuestro alrededor la elección puede ser mucho más rápida. La actitud hacia las conjeturas es bastante semejante. Siendo iguales en otras cosas, una conjetura que tiene muchas rivales evidentes resulta más difícil de aceptar que la que carece de rivales. Si usted piensa igual que yo, no habrá duda que en este aspecto la ventaja cae sobre nuestra conjetura, y no sobre la de Bachet.

En fin, observemos que la evidencia para la conjetura de Bachet es más fuerte en un aspecto, mientras la evidencia para la nuestra es más fuerte en otros aspectos; y dejemos de plantearnos problemas incontestables.

## EJEMPLOS Y COMENTARIOS

**1.** *Notación.* Supongamos que $n$ y $k$ son números enteros positivos y consideremos la ecuación diofantina

$$n = x_1^2 + x_2^2 + \ldots + x_k^2.$$

Decimos que dos soluciones $x_1, x_2, \ldots x_k$ y $x'_1\, x'_2, \ldots x'_k$ son iguales, si, y sólo si, $x_1 = x'_1$, $x_2 = x'_2$, $\ldots$ $x_k = x_k$. Si admitimos para $x_1, x_2, \ldots x_k$ números *todos* enteros, positivos, negativos o nulos, llamaremos al número de soluciones $R_k(n)$. Si admitimos sólo enteros *impares positivos*, llamaremos al número de soluciones $S_k(n)$. Esta notación es importante en la mayoría de los problemas siguientes.

La conjetura de Bachet (sección 2) es expresada en esta notación por la desigualdad

$$R_4(n) > 0, \quad \text{para } n = 1, 2, 3, \ldots$$

La conjetura que nosotros descubrimos en la sección 6 afirma que $S_4(4(2n - 1))$ es igual a la suma de los divisores de $2n - 1$, para $n = 1, 2, 3, \ldots$

Encontrar $R_2(25)$ y $S_3(11)$.

**2.** Sean $x$ e $y$ coordenadas rectangulares en un plano. Los puntos para lo que $x$ e $y$ son enteros se llaman «puntos de enlace» del plano. Los puntos de enlace del espacio son definidos de modo semejante.

Interpretar $R_2(n)$ y $R_3(n)$ geométricamente, en términos de puntos de enlace.

**3.** Expresar la conjetura encontrada en la sección 1, usando el símbolo $R_2(n)$.

**4.** ¿Cuándo es un número primo impar la suma de dos cuadrados? Intentar contestar esta cuestión inductivamente, mediante el examen de la tabla

$$\begin{array}{rcl} 3 & & - \\ 5 & = & 4 + 1 \\ 7 & & - \\ 11 & & - \\ 13 & = & 9 + 4 \\ 17 & = & 16 + 1 \\ 19 & & - \\ 23 & & - \\ 29 & = & 25 + 4 \\ 31 & & - \end{array}$$

Ampliar esta tabla si es necesario y compararla con la tabla de la sección 1.

**5.** ¿Podría verificar por deducción matemática alguna parte de su contestación al ejemplo 4 obtenida por inducción? Después de tal verificación, ¿sería razonable cambiar nuestra confianza en la conjetura?

**6.** Verificar la conjetura de Bachet (sección 2) hasta 30 inclusive. ¿Qué números requieren en realidad cuatro cuadrados?

**7.** Para comprender mejor la tabla I de la sección 5 denotemos $a^2$, $b^2$, $c^2$ y $d^2$, cuatro cuadrados impares diferentes y consideremos las sumas

$$\begin{aligned}
&(1)\quad a^2+b^2+c^2+d^2\\
&(2)\quad a^2+a^2+b^2+c^2\\
&(3)\quad a^2+a^2+b^2+b^2\\
&(4)\quad a^2+a^2+a^2+b^2\\
&(5)\quad a^2+a^2+a^2+a^2.
\end{aligned}$$

¿Cuántas representaciones diferentes (en el sentido de la sección 3) se pueden derivar de cada una por permutación de los términos?

**8.** El número de representaciones de $4u$, como una suma de cuatro cuadrados impares es impar si, y sólo si, $u$ es un cuadrado. (Siguiendo la notación de la sección 3, consideramos que $u$ es impar.) Probar esta proposición y mostrar que se halla de acuerdo con la conjetura de la sección 6. ¿Cómo influye esta observación sobre nuestra confianza en la conjetura?

**9.** Ahora, denotemos con $a$, $b$, $c$ y $d$ diferentes enteros positivos (pares e impares). Consideremos las cinco sumas mencionadas en el ejemplo 7 y estas otras:

$$\begin{aligned}
&(6)\quad a^2+b^2+c^2 \qquad &&(9)\quad a^2+b^2\\
&(7)\quad a^2+a^2+b^2 \qquad &&(10)\quad a^2+a^2\\
&(8)\quad a^2+a^2+a^2 \qquad &&(11)\quad a^2.
\end{aligned}$$

Encontrar en cada uno de estos once casos la contribución a $R_4(n)$. Derivar de cada suma todas las representaciones posibles mediante las siguientes operaciones obvias: añadir $0^2$ tantas veces como sea necesario para que el número de términos llegue a cuatro, cambiar su ordenación y remplazar algunos (o ninguno, o todos) de los números $a$, $b$, $c$, $d$ por $-a$, $-b$, $-c$, $-d$, respectivamente. (Comprobar los ejemplos en la tabla II.)

**10.** Investigar inductivamente el número de soluciones de la ecuación $n = x^2+y^2+z^2+w^2$ en enteros, $x$, $y$, $z$ y $w$, positivos, negativos, o 0. Empecemos construyendo una tabla análoga a la tabla I.

**11.** (Continúa.) Intentar usar el método o el resultado de la sección 6.

**12.** (Continúa.) Guiados por la analogía de la sección 6 o por observación de la tabla II, distinguir las clases apropiadas de enteros e investigar cada una por sí misma.

**13.** (Continúa.) Concentrémonos sobre las clases más rebeldes.

**14.** (Continúa.) Intentar resumir todas las regularidades fragmentarias y expresar la ley en una frase.

**15.** (Continúa.) Comprobar la regla encontrada en los primeros tres casos no contenidos en la tabla II.

**16.** Encontrar $R_8$ (5) y $S_8$ (40).

**17.** Comprobar por lo menos dos entradas (notaciones) de la tabla III, página 113, aún no dadas en las tablas I y II.

**18.** Investigar inductivamente $R_8$ ($n$) y $S_8$, usando la tabla III.

**19.** (Continúa.) Intentar usar el método, o el resultado, de la sección 6 y el ejemplo 10-15.

**20.** (Continúa.) Mediante la analogía y la observación, distinguir las clases adecuadas de enteros e investigar cada una por sí misma.

**21.** (Continúa.) Intentar extraer una clave del caso más accesible.

**22.** (Continúa.) Encontrar un concepto unificante que pueda resumir las regularidades fragmentarias.

**23.** (Continúa.) Intentar expresar la ley en una frase.

**24.** ¿Qué números pueden y cuáles no pueden ser expresados por la forma $3x + 5y$, donde $x$ e $y$ son enteros no-negativos?

**25.** Intentar intuir la ley de la siguiente tabla:

| $a$ | $b$ | Ultimo entero no expresable por la forma $ax + by$ |
|---|---|---|
| 2 | 3 | 1 |
| 2 | 5 | 3 |
| 2 | 7 | 5 |
| 2 | 9 | 7 |
| 3 | 4 | 5 |
| 3 | 5 | 7 |
| 3 | 7 | 11 |
| 3 | 8 | 13 |
| 4 | 5 | 11 |
| 5 | 6 | 19 |

Se comprende que $x$ e $y$ son enteros no-negativos Comprobemos ahora algunos puntos más, si es necesario. [Observar el cambio en la última columna cuando sólo uno de los dos números, $a$ y $b$, cambia.]

**26.** *Peligros de inducción.* Examinemos inductivamente las siguientes aserciones:

(1) $(n-1)!+1$ es divisible por $n$ cuando $n$ es un primo, pero no divisible por $n$ cuando éste es compuesto.

(2) $2^{n-1}-1$ es divisible por $n$ cuando $n$ es un primo impar, pero no divisible por $n$ cuando $n$ es compuesto.

TABLA II

| $n$ | No incrementan | Representaciones | $R_4(n)/8$ |
|---|---|---|---|
| 1 | 1 | $4 \times 2$ | 1 |
| 2 | $1+1$ | $6 \times 4$ | 3 |
| 3 | $1+1+1$ | $4 \times 8$ | 4 |
| 4 | 4 | $4 \times 2$ | 3 |
| | $1+1+1+1$ | $1 \times 16$ | |
| 5 | $4+1$ | $12 \times 4$ | 6 |
| 6 | $4+1+1$ | $12 \times 8$ | 12 |
| 7 | $4+1+1+1$ | $4 \times 16$ | 8 |
| 8 | $4+4$ | $6 \times 4$ | 3 |
| 9 | 9 | $4 \times 2$ | 13 |
| | $4+4+1$ | $12 \times 8$ | |
| 10 | $9+1$ | $12 \times 4$ | 18 |
| | $4+4+1+1$ | $6 \times 16$ | |
| 11 | $9+1+1$ | $12 \times 8$ | 12 |
| 12 | $9+1+1+1$ | $4 \times 16$ | 12 |
| | $4+4+4$ | $4 \times 8$ | |
| 13 | $9+4$ | $12 \times 4$ | 14 |
| | $4+4+4+1$ | $4 \times 16$ | |
| 14 | $9+4+1$ | $24 \times 8$ | 24 |
| 15 | $9+4+1+1$ | $12 \times 16$ | 24 |
| 16 | 16 | $4 \times 2$ | 3 |
| | $4+4+4+4$ | $1 \times 16$ | |
| 17 | $16+1$ | $12 \times 4$ | 18 |
| | $9+4+4$ | $12 \times 8$ | |
| 18 | $16+1+1$ | $12 \times 8$ | 39 |
| | $9+9$ | $6 \times 4$ | |
| | $9+4+4+1$ | $12 \times 16$ | |
| 19 | $16+1+1+1$ | $4 \times 16$ | 20 |
| | $9+9+1$ | $12 \times 8$ | |
| 20 | $16+4$ | $12 \times 4$ | 18 |
| | $9+9+1+1$ | $6 \times 16$ | |
| 21 | $16+4+1$ | $24 \times 8$ | 32 |
| | $9+4+4+4$ | $4 \times 16$ | |
| 22 | $16+4+1+1$ | $12 \times 16$ | 36 |
| | $9+9+4$ | $12 \times 8$ | |

TABLA II (*Continuación*)

| $n$ | No incrementa | Representación | $R_4(n)/8$ |
|---|---|---|---|
| 23 | 9 + 9 + 4 + 1 | 12 × 16 | 24 |
| 24 | 16 + 4 + 4 | 12 + 8 | 12 |
| 25 | 25 | 4 × 2 | 31 |
| | 16 + 9 | 12 × 4 | |
| | 16 + 4 + 4 + 1 | 12 × 16 | |
| 26 | 25 + 1 | 12 × 4 | 42 |
| | 16 + 9 + 1 | 24 × 8 | |
| | 9 + 9 + 4 + 4 | 6 × 16 | |
| 27 | 25 + 1 + 1 | 12 × 8 | 40 |
| | 16 + 9 + 1 + 1 | 12 × 16 | |
| | 9 + 9 + 9 | 4 × 8 | |
| 28 | 25 + 1 + 1 + 1 | 4 × 16 | 24 |
| | 16 + 4 + 4 + 4 | 4 × 16 | |
| | 9 + 9 + 9 + 1 | 4 × 16 | |
| 29 | 25 + 4 | 12 × 4 | 30 |
| | 16 + 9 + 4 | 24 × 8 | |
| 30 | 25 + 4 + 1 | 24 × 8 | 72 |
| | 16 + 9 + 4 + 1 | 24 × 16 | |

TABLA III

| $n$ | $R_4(n)/8$ | $R_8(n)/16$ | $S_8(8n)$ | $S_4(4(2n-1))$ | $2n-1$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 3 | 7 | 8 | 4 | 3 |
| 3 | 4 | 28 | 28 | 6 | 5 |
| 4 | 3 | 71 | 64 | 8 | 7 |
| 5 | 6 | 126 | 126 | 13 | 9 |
| 6 | 12 | 196 | 224 | 12 | 11 |
| 7 | 8 | 344 | 344 | 14 | 13 |
| 8 | 3 | 583 | 512 | 24 | 15 |
| 9 | 13 | 757 | 757 | 18 | 17 |
| 10 | 18 | 882 | 1008 | 20 | 19 |
| 11 | 12 | 1332 | 1332 | 32 | 21 |
| 12 | 12 | 1988 | 1792 | 24 | 23 |
| 13 | 14 | 2198 | 2198 | 31 | 25 |
| 14 | 24 | 2408 | 2752 | 40 | 27 |
| 15 | 24 | 3528 | 3528 | 30 | 29 |
| 16 | 3 | 4679 | 4096 | 32 | 31 |
| 17 | 18 | 4914 | 4914 | 48 | 33 |
| 18 | 39 | 5299 | 6056 | 48 | 35 |
| 19 | 20 | 6860 | 6860 | 38 | 37 |
| 20 | 18 | 8949 | 8064 | 56 | 39 |

# Ejemplos variados de inducción

*Cuando usted se ha convencido a sí mismo de que el teorema es cierto debe empezar a demostrarlo.*—EL PROFESOR TRADICIONAL DE MATEMÁTICAS *.

## 1. DESARROLLOS

Cuando nos enfrentamos con cualquier clase de problemas es raro que no necesitemos de algún tipo de razonamiento inductivo. En algunas ramas de las matemáticas, se requiere el uso del razonamiento inductivo de una manera característica. Este capítulo trata de ilustrar este punto con algunos ejemplos. Empezaremos con uno relativamente sencillo.

*Desarróllense los valores de x en la función* $1/(1 - x + x^2)$.

Este problema puede ser resuelto de varias maneras. La siguiente solución es algo tosca, pero está basada sobre un conocido principio y puede ocurrírsele a un principiante que sabe poco, con tal que conozca la suma de las series geométricas:

$$1 + r + r^2 + r^3 + \ldots = \frac{1}{1 - r}.$$

Hay oportunidad de usar esta fórmula en nuestro problema:

$$\frac{1}{1 - x + x^2} = \frac{1}{1 - x(1 - x)}$$

$$= 1 + x(1 - x) + x^2(1 - x)^2 + x^3(1 - x)^3 + \ldots$$

$$\begin{aligned}
= 1 &+ x - x^2 \\
&+ x^2 - 2x^3 + x^4 \\
&+ x^3 - 3x^4 + 3x^5 - x^6 \\
&+ x^4 - 4x^5 + 6x^6 - 4x^7 + x^8 \\
&+ x^5 - 5x^6 + 10x^7 - 10x^8 + \ldots \\
&+ x^6 - 6x^7 + 15x^8 - \ldots \\
&+ x^7 - 7x^8 + \ldots \\
&+ x^8 - \ldots \\
&\ldots \\
= 1 + x &- x^3 - x^4 + x^6 + x^7 \qquad \ldots
\end{aligned}$$

* Este aforismo del conocido pedagogo (*How to Solve It*, pág. 181) es precedido algunas veces de esta exhortación: «Si tiene que demostrar un teorema no se precipite. En primer lugar, comprenda completamente lo que el teorema dice, intente ver clara-

El resultado es chocante. Ningún coeficiente que no desaparezca tiene el valor 1 ó —1. La sucesión de los coeficientes parece mostrar alguna regularidad, que se hace más manifiesta si calculamos más términos:

$$\frac{1}{1-x+x^2} = 1 + x - x^3 - x^4 + x^6 + x^7 - x^9 - x^{10} + x^{12} + x^{13} - \ldots$$

¡Periódica! La secuencia de los coeficientes aparece como periódica con el período 6:

$$1, 1, 0, -1, -1, 0 \mid 1, 1, 0, -1, -1, 0 \mid 1, 1, \ldots$$

Naturalmente, esperamos que la periodicidad observada se extienda más allá de los límites de nuestra observación. Sin embargo, esto es una conclusión inductiva, o una mera intuición, que miraremos con justificado escepticismo. La intuición, no obstante, está basada sobre los hechos, y por ello merece serio examen. Examinarla significa, entre otras cosas, apoyarla. Hay una manera interesante de apoyar nuestra conjetura:

$$\frac{1}{1-x+x^2} = 1 - x^3 + x^6 - x^9 + x^{12} - \ldots$$
$$+ x - x^4 + x^7 - x^{10} + x^{13} - \ldots$$

Ahora podemos observar fácilmente dos series geométricas al lado derecho, ambas con la misma razón $-x^3$, que podemos sumar. Y así, nuestra conjetura se reduce a

$$\frac{1}{1-x+x^2} = \frac{1}{1+x^3} + \frac{x}{1+x^3} = \frac{1+x}{1+x^3},$$

que es, por supuesto, cierta. Hemos probado la conjetura.

Nuestro ejemplo, aun siendo tan sencillo, es típico en varios aspectos. Si tenemos que desarrollar una función dada podemos obtener los primeros coeficientes sin mucho trabajo. Mirando esos coeficientes intentaremos, como hemos hecho aquí, intuir la ley que gobierna el desarrollo. Una vez intuida ésta trataremos de probarla. Puede ser una gran ventaja, como lo ha sido en este caso, llevar la prueba hacia atrás, empezando con una enunciación de la conjetura clara y adecuada.

---

mente lo que significa. Luego, compruébelo; puede ser falso. Examine sus consecuencias, verifíquelo en varios ejemplos particulares hasta convencerse a sí mismo de que es cierto. Cuando...».

Por cierto que nuestro ejemplo es bastante satisfactorio (además de típico), pues nos conduce a una curiosa relación entre coeficientes de binomios.

No es superfluo añadir que el problema de desarrollar una función dada en una serie surge frecuentemente en diversas ramas de la matemática. Véase la próxima sección y los ejemplos y comentarios del capítulo VI.

## 2. Aproximaciones *

Denotemos con $E$ la longitud del perímetro de una elipse con semiejes $a$ y $b$. No hay una expresión sencilla para $E$ en términos de $a$ y $b$, pero varias expresiones aproximadas se han propuesto, entre las cuales las dos siguientes son quizá las más obvias:

$$P = \pi (a + b), \qquad P' = 2\pi (ab)^{1/2};$$

$P$ y $P'$ son expresiones aproximadas de la longitud del perímetro de la elipse, y $E$ es la expresión exacta de esa misma cantidad. Cuando $a$ coincide con $b$ la elipse se convierte en un círculo y ambos $P$ y $P'$ coinciden con $E$.

¿Cuánto se aproximan $P$ y $P'$ a $E$, cuando $a$ es diferente de $b$? ¿Cuál se acerca más a la verdad, $P$ o $P'$? Cuestiones de este tipo se presentan con frecuencia en todas las ramas de la matemática aplicada, donde existe un procedimiento ampliamente aceptado para resolverlas, y que puede ser descrito a grandes rasgos como sigue: *Desarrollen* $(P—E)/E$, *el error relativo de aproximación, en valores de una pequeña cantidad adecuada y fundamenten el juicio sobre el término inicial* (el primero que no desaparece) *del desarrollo.*

Veamos ahora lo que esto significa y cómo funciona el procedimiento cuando lo aplicamos a nuestro caso. Primero, elijamos una «cantidad pequeña adecuada». Probemos con $\varepsilon$, la excentricidad numérica de la elipse, definida por esta fórmula:

$$\varepsilon = \frac{(a^2 - b^2)^{1/2}}{a};$$

tomemos $a$ como el semieje mayor y $b$ como el menor. Cuando $a$ coincide con $b$ la elipse se convierte en círculo y $\varepsilon$ desaparece. Cuando la elipse no es muy diferente de un círculo, $\varepsilon$ es muy pequeña. Por tanto, des-

* Cf. Putnam, 1949.

arrollemos el error relativo en los valores de $\varepsilon$. Obtenemos (omitimos los detalles aquí) entonces

$$\frac{P - E}{E} = -\frac{1}{64}\varepsilon^4 + \ldots, \quad \frac{P' - E}{E} = -\frac{3}{64}\varepsilon^4 + \ldots$$

Calculamos el error inicial que en ambos casos, siendo 4, contiene $\varepsilon^4$. Hemos omitido en ambos desarrollos los términos más altos, que contienen $\varepsilon^5$, $\varepsilon^6$, ... Los términos omitidos son despreciables cuando $\varepsilon$ es muy pequeña (infinitamente pequeña), es decir, cuando la elipse es casi circular. Por tanto, para elipses casi circulares, $P$ está más próximo al valor verdadero que $P'$. (En efecto, la razón de errores alcanza 1 : 3 cuando $\varepsilon$ tiende a 0.) Ambos $P$ y $P'$ se aproximan a $E$ así:

$$E > P > P'.$$

Todo esto es válido para $\varepsilon$ muy pequeñas, para elipses casi circulares Sin embargo, aún no sabemos cuántos de estos resultados son válidos cuando $\varepsilon$ no es tan pequeña. De hecho, ahora sólo sabemos relaciones límites para $\varepsilon \longrightarrow 0$. No sabemos nada definitivo sobre el error de nuestras aproximaciones cuando $\varepsilon = 0.5$ o $\varepsilon = 0.1$. Por supuesto, lo que necesitamos en la práctica es información sobre estos casos concretos.

En tales circunstancias, las gentes prácticas ensayan sus fórmulas numéricamente. En esto podríamos imitarles. Pero, ¿qué caso ensayaremos primero? Es aconsejable no olvidar los casos extremos. La excentricidad numérica $\varepsilon$ varía entre los valores extremos 0 y 1. Cuando $\varepsilon = 0$, $b = a$ y la elipse se convierte en círculo. Ahora conocemos muy bien este caso; veamos el extremo. Cuando $\varepsilon = 1$, $b = 0$, la elipse se convierte en un segmento de longitud $2a$ y la longitud del perímetro será $4a$. Tenemos, pues,

$$E = 4a, \quad P = \pi a, \quad P' = 0 \text{ cuando } \varepsilon = 1.$$

Puede ser interesante notar que, en ambos casos, para $\varepsilon = 1$ tanto como para el más pequeño $\varepsilon$, $E > P > P'$. ¿Son estas desigualdades generalmente válidas?

Para la segunda desigualdad la contestación es fácil. En efecto, tenemos para $a > b$,

$$P = \pi (a + b) > 2\pi (ab)^{1/2} = P',$$

pues esto es equivalente a

$$(a + b)^2 > 4ab,$$

o a

$$(a - b)^2 > 0.$$

Dirijamos la atención sobre la otra cuestión. ¿Es la desigualdad $E > P$ generalmente válida? Es natural conjeturar que lo que encontramos cierto en casos extremos ($\varepsilon$ pequeña, $\varepsilon = 1$) será igualmente cierto en los casos intermedios (para todos los valores de $\varepsilon$ entre 0 y 1). Nuestra conjetura no está apoyada en muchas observaciones, es verdad, pero la analogía le da valor. Una cuestión parecida (concerniente a $P > P'$) que nos preguntamos en el mismo tono y con un fundamento semejante fue contestada en sentido afirmativo.

Probemos, con un caso, numéricamente. Sabemos un poco más del caso en que $\varepsilon$ está cerca de 0 que del caso en que está cerca de 1. Elegimos un valor simple para $\varepsilon$ más cerca de 1 que de 0: $a = 5$, $b = 3$, $\varepsilon = 4/5$. Encontramos para esta $\varepsilon$ (usando tablas apropiadas) que

$$E = 2\pi \times 4.06275, \qquad P = 2\pi \times 4.00000.$$

La desigualdad $E > P$ está verificada. Esta verificación numérica de nuestra conjetura viene de un nuevo lado, de una fuente diferente, y por ello es de cierto peso. Notemos también que

$$(P - E)/E = -0.0155, \qquad -\varepsilon^4/64 = -0.0064.$$

El error relativo es sobre 1,5 por 100. Es considerablemente mayor que el término inicial de su desarrollo, pero tiene el mismo signo. Como $\varepsilon = 4/5 = 0.8$ no es demasiado pequeño, nuestra observación se adecua al conjunto del cuadro y tiende a aumentar la confianza en la conjetura.

Las fórmulas de aproximación juegan un papel importante en la matemática aplicada. Intentando juzgar una fórmula tal, muy a menudo adoptamos en la práctica el procedimiento seguido en esta sección. Calculamos el término inicial en el desarrollo del error relativo y luego complementamos la información así adquirida con pruebas numéricas, consideraciones de analogía, etc.; en una palabra, con razonamiento inductivo y no-demostrativo.

## 3. Límites

Con objeto de ver cómo funciona el razonamiento inductivo en otro dominio, consideremos el siguiente problema *.

*Sea* $a_1, a_2, \ldots, a_n, \ldots,$ *una secuencia arbitraria de números positivos. Mostrar que*

$$\limsup_{n \to \infty} \left( \frac{a_1 + a_{n+1}}{a_n} \right)^n \geq e.$$

* Ver Putnam, 1948.

Este problema requiere algún conocimiento preliminar, especialmente estar familiarizado con el concepto de «lím sup» o «límite más alto de indeterminación»* Aun los familiarizados con este concepto experimentarán dificultades en encontrar una prueba. Mi enhorabuena a todos los estudiantes que logren resolverlo por sus propios medios en pocas horas.

Si usted ha luchado por su cuenta con el problema antes, seguirá con más simpatía el intento descrito en las páginas siguientes.

## 4. Intentando refutarlo

Empecemos con las preguntas matemáticas usuales.

*¿Cuál es la hipótesis?* Simplemente, que $a_n > 0$.

*¿Cuál es la conclusión?* La desigualdad con $e$ a la derecha y el complicado límite de la izquierda.

*¿Conoce un teorema relacionado?* No, ciertamente. Es muy diferente de todo lo que conozco.

*¿Es verosímil que el teorema sea cierto? ¿O es más verosímil que sea falso?* Falso, por supuesto. Es difícil creer que una consecuencia tan precisa pueda derivarse de la simple hipótesis de que $a_n > 0$

*¿Qué es lo que se pide?* Probar el teorema o refutarlo. Por mi parte, me inclino más a lo último.

*¿Podemos probar algún caso particular del teorema?* Eso es lo primero que vamos a hacer.

[Con objeto de simplificar las fórmulas proponemos

$$\left(\frac{a_1 + a_{n+1}}{a_n}\right)^n = b_n$$

y consideramos $b_n \to b$ para $\lim_{n\to\infty} b_n = b$.]

Si hacemos $a_n = 1$, para $n = 1, 2, 3, \ldots$, entonces

$$b_n = \left(\frac{1 + 1}{1}\right)^n = 2^n \to \infty .$$

En este caso, el enunciado del teorema se verifica.

Ahora propongamos $a_1 = 0$, $a_n = 1$ para $n = 2, 3, 4, \ldots$, de donde

$$b^n = \left(\frac{0 + 1}{1}\right)^n = 1^n \to 1 < e.$$

---

* Ver G. H. Hardy, *Pure Mathematics*, sec. 82.

¡El teorema queda desbaratado! No, no es eso. La hipótesis permite que sea $a_1 = 0.00001$, pero no $a_1 = 0$. ¡Qué lástima!

Intentemos algo parecido. Sea $a_n = n$, de donde

$$b_n = \left(\frac{1 + (n + 1)}{n}\right)^n = \left(1 + \frac{2}{n}\right)^n \to e^2.$$

De nuevo queda verificado.

Ahora tomemos $a_n = n^2$, de donde

$$b_n = \left(\frac{1 + (n + 1)^2}{n^2}\right)^n = \left(1 + \frac{2}{n}\,\frac{n + 1}{n}\right)^n \to e^2.$$

Otra vez verificado. Y otra vez $e^2$. ¿Debemos colocar $e^2$ en la conclusión del lado derecho en lugar de $e$? Eso mejoraría el teorema.

*Introduzcamos un parámetro.* Tomemos... Sí, tomemos $a_1 = c$, donde $c$ es un elemento del que podemos disponer, aunque $a_n = n$ para $n = 2, 3, 4, \ldots$ De aquí

$$b_n = \left(\frac{c + (n + 1)}{n}\right)^n = \left(1 + \frac{1 + c}{n}\right)^n \to e^{1+c}.$$

Esto es siempre $> e$, puesto que $c = a_1 > 0$, aun cuando podemos acercarlo a $e$ tanto como queramos, ya que $c$ puede ser arbitrariamente empequeñecido. Por ahora, resulta imposible refutarlo o demostrarlo.

Hagamos todavía un nuevo ensayo. Tomemos $a_n = n^c$. Luego [omitidos algunos cálculos],

$$b_n = \left[\frac{1 + (n + 1)^c}{n^c}\right]^n \to \begin{cases} \infty & \text{Si } 0 < c < 1, \\ e^2 & \text{Si } c = 1, \\ e^c & \text{Si } c > 1. \end{cases}$$

De nuevo el límite puede acercarse a $e$ tanto cuanto queramos, aunque siempre permanezca superior a $e$. Nunca lograremos traerlo por debajo de este... límite bajo $e$. Demos, pues, la vuelta.

## 5. Intentando probarlo.

En efecto, las indicaciones para considerar esta nueva posibilidad son bastante fuertes. A la luz de la evidencia inductiva acumulada, las perspectivas de refutar el teorema aparecen tan disminuidas que las perspectivas de probarlo se hacen algo más vivas.

En consecuencia, no queda otra cosa que volver a examinar el teorema, su enunciación, su hipótesis, su conclusión, los conceptos que implica, etc.

*¿Podemos mitigar la hipótesis?* No, no podemos. Si admitimos que $a_n = 0$, la conclusión no será nunca válida y el teorema es falso ($a_1 = 0$, $a_2 = a_3 = a_4 = \ldots = 1$).

*¿Podemos mejorar la conclusión?* En realidad, no podemos mejorarla sustituyendo $e$ por algún número mayor, pues entonces la conclusión no sería válida y el teorema se haría falso (ver ejemplos en la anterior sección 4).

*¿Hemos tomado en cuenta todas las nociones esenciales envueltas en el problema?* No; y aquí puede estar el error.

*¿Qué es lo que no hemos tenido en cuenta?* La definición de lím sup y la definición del número $e$.

¿Qué es lím sup $b_n$? Es el límite superior de indeterminación para $b_n$ cuando $n$ tiende a infinito.

¿Qué es $e$? Puede ser definida de diversos modos, pero los anteriores ejemplos sugieren que la mejor definición será la más familiar:

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

*¿Podríamos formular de nuevo el teorema?* . . . . . .
*¿Podríamos formularlo de alguna manera más accesible?* . . . . . .
*¿Podríamos formular de nuevo la conclusión? ¿Cuál es la conclusión?* La conclusión contiene $e$. ¿Qué es $e$? (Antes fracasamos al contestar esto.) ¡Oh, sí!; la conclusión es

$$\lim_{n\to\infty} \sup \left(\frac{a_1 + a_{n+1}}{a_n}\right)^n \geqq \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

o, lo que es lo mismo,

$$\lim_{n\to\infty} \sup \left[\frac{n\,(a_1 + a_{n+1})}{(n+1)\,a_n}\right]^n \geqq 1.$$

¡Esto parece mucho mejor!

*¿Puede ser falsa la conclusión, cuando se cumple la hipótesis?* Sí, ésa es la cuestión. Veamos. Por lo pronto, examinemos la negación del enunciado, esto es, el enunciado opuesto. Escribámoslo:

$$(?) \qquad \lim_{n\to\infty} \sup \left[\frac{n(a_1 + a_{n+1})}{(n+1)a_n}\right]^n < 1.$$

Ponemos una interrogación al lado, precisamente porque éste es un punto en duda. Llamémosla «la fórmula (?)». ¿Qué significa (?)? En realidad, implica que hay una $N$ tal que

$$\left[\frac{n\,(a_1 + a_{n+1})}{(n+1)\,a_n}\right]^n < 1 \text{ para } n \geqq N.$$

De donde se sigue que

$$\frac{n\,(a_1 + a_{n+1})}{(n+1)a_n} < 1 \text{ para } n \geqq N.$$

De donde se sigue ... Intentemos algo nuevo. ¡Escribamos con nitidez! Se sigue, además de (?), que

$$\frac{a_1}{n+1} + \frac{a_{n+1}}{n+1} < \frac{a_n}{n},$$

o,

$$\frac{a_{n+1}}{n+1} - \frac{a_n}{n} < -\frac{a_1}{n+1} \text{ para } n \geqq N.$$

Escribámoslo ampliamente. Se sigue de eso

$$\frac{a_n}{n} - \frac{a_{n-1}}{n-1} < -\frac{a_1}{n}$$

$$\frac{a_{n-1}}{n-1} - \frac{a_{n-2}}{n-2} < -\frac{a_1}{n-1}$$

. . . . . .

$$\frac{a_{N+1}}{a_{N+1}} - \frac{a_N}{N} < -\frac{a_1}{N+1}$$

y, por tanto,

$$\frac{a_n}{n} < \frac{a_N}{N} - a_1\left(\frac{1}{N+1} + \frac{1}{N+2} + \cdots + \frac{1}{n-1} + \frac{1}{n}\right)$$

$$= C - a_1\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right),$$

donde $C$ es una constante, independiente de $n$, habida cuenta que $n \geq N$. Esto no importa realmente, pero, de hecho,

$$C = \frac{a_N}{N} + a_1 \left(1 + \frac{1}{2} + \ldots + \frac{1}{N}\right).$$

Esto supone que $n$ puede ser arbitrariamente grande y que la serie armónica se hace divergente. De aquí se sigue, por tanto, que

$$\lim_{n \to \infty} \frac{a_n}{n} = -\infty .$$

Ahora bien, esto contradice llanamente la hipótesis de que $a_n > 0$ para $n = 1, 2, 3, \ldots$ Sin embargo, esta contradicción se sigue sin falta de la fórmula (?). En consecuencia, (?) es de hecho responsable de la contradicción; (?) es incompatible con la hipótesis $a_n > 0$; la opuesta a (?) tiene, pues, que ser cierta; ¡el teorema está probado!

## 6. El papel de la fase inductiva

Mirando superficialmente la solución precedente estamos tentados de pensar que la primera fase inductiva, la de la solución (sección 4) no es usada absolutamente en la segunda fase, la demostrativa (sección 5). Sin embargo, esto no es así. La fase inductiva fue útil en varios aspectos.

Primero, examinando casos concretos y particulares del teorema, llegamos a comprenderlo, a darnos cuenta de su significado. Nos satisfizo que su hipótesis es esencial y su conclusión sagaz. Esta información nos fue muy valiosa en la segunda parte: supimos que debíamos utilizar la hipótesis total y que debíamos tener en cuenta el valor preciso de la constante $e$.

Segundo, habiendo verificado el teorema en varios casos particulares, sacamos una fuerte evidencia inductiva de ellos. La fase inductiva superó nuestra sospecha inicial y nos dio mucha confianza en el teorema. Sin tal confianza no hubiéramos tenido ánimo de emprender la prueba, que no parecía, ni mucho menos, un trabajo rutinario. «Cuando usted se ha convencido a sí mismo de que el teorema es cierto debe empezar a demostrarlo»; el profesor tradicional de matemáticas tiene toda la razón.

Tercero, los ejemplos en que la fórmula familiar del límite para $e$

nos sorprendieron una y otra vez, nos dio pie para introducir razonablemente esa fórmula del límite en el enunciado del teorema. Y esa introducción fue un paso crucial para la solución.

En conjunto, parece razonable y natural que la fase inductiva preceda a la fase demostrativa. Primero, intuir; luego, probar.

## EJEMPLOS Y COMENTARIOS

**1.** Multiplicando las series

$$(1 - x^2)^{-1/2} = 1 + \frac{1}{2}x^2 + \frac{1}{2}\frac{3}{4}x^4 + \dots$$

$$\text{arc sen } x = \frac{x}{1} + \frac{1}{2}\frac{x^3}{3} + \frac{1}{2}\frac{3}{4}\frac{x^5}{5} + \dots$$

usted encontrará los primeros términos del desarrollo

$$y = (1 - x^2)^{-1/2} \text{ arc sen } x = x + \frac{2}{3}x^3 + \dots$$

(*a*) Calcúlense algunos términos más y trátase de intuir el término general.

(*b*) Pruébese que $y$ satisface la ecuación diferencial

$$(1 - x^2)\,y' - xy = 1$$

y usar esta ecuación para probar la intuición anterior.

**2.** Multiplicando las series

$$e^{x^2/2} = 1 + \frac{x^2}{2} + \frac{x^4}{2 \cdot 4} + \dots$$

$$\int_0^x e^{-t^2/2}dt = \frac{x}{1} - \frac{1}{2}\frac{x^3}{3} + \frac{1}{2 \cdot 4}\frac{x^5}{5} - \dots$$

encontraremos los primeros términos del desarrollo

$$y = e^{x^2/2}\int_0^x e^{-t^2/2}dt = x + \frac{1}{3}x^3 + \dots$$

(*a*) Calcular algunos términos más y tratar de intuir el término general.

(*b*) Esta intuición, si es correcta, sugiere que $y$ satisface una ecuación diferencial sencilla. Probar la intuición mediante el establecimiento de la ecuación.

**3.** La ecuación funcional

$$f(x) = \frac{1}{1+x} f\left(\frac{2\sqrt{x}}{1+x}\right)$$

se satisface con los valores de la serie

$$f(x) = 1 + \frac{1}{4}x^2 + \frac{9}{64}x^4 + \frac{25}{256}x^6 + \frac{1225}{16384}x^8 + \ldots$$

Verificar estos coeficientes, derivar algunos más, si es necesario, e intentar intuir el término general.

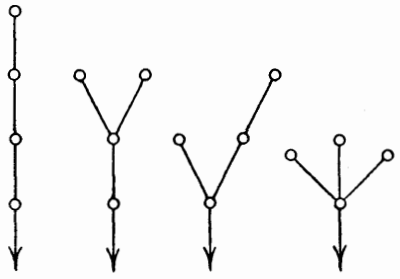


Fig. 5.1. Compuestos $C_4H_9OH$.

**4.** La ecuación funcional

$$f(x) = 1 + \frac{x}{6}\left[f(x)^3 + 3f(x)f(x^2) + 2f(x^3)\right]$$

se satisface con los valores de la serie

$$f(x) = 1 + x + x^2 + 2x^3 + 4x^4 + \ldots + a_n x^n + \ldots$$

Se asegura que $a_n$ es el número de los compuestos químicos estructuralmente diferentes (alcoholes alifáticos), que tienen la misma fórmula química $C_nH_{2n+1}OH$. En el caso $n = 4$, la contestación es verdadera. Hay $a_4 = 4$ alcoholes $C_4H_9OH$; los encontraremos representados en la figura 5.1, donde cada compuesto tiene un «árbol», cada $C$ es un pequeño

círculo o «nudo» y el radical — OH es una flecha; las $H$ han caído. Probar con otros valores de $n$.

**5.** $$\sum_{k=0}^{n/2} (-1)^k \left(\frac{n-k}{k}\right) = 1, 1, 0, -1, -1, 0,$$

según $n \equiv 0, 1, 2, 3, 4, 5$ (modelo 6).

**6.** Una elipse describe un esferoide alargado o achatado, según que gire sobre su eje mayor o menor.

Para el área de la superficie del esferoide alargado

$$E = 2\pi ab\,[(1 - \varepsilon^2)^{1/2} + (\text{arc sen } \varepsilon)/\varepsilon],\ P = 4\pi\,(a^2 + 2\,b^2)/3$$

son las expresiones exacta y aproximada, respectivamente ($a$, $b$ y $\varepsilon$ como en la sección 2). Encontrar

($a$) el término inicial del error relativo,

($b$) el error relativo cuando $b = 0$.

¿Qué pasa con el signo del error relativo?

**7.** Para el área de la superficie del elipsoide achatado

$$E = 2\pi a^2 \left[1 + \frac{1-\varepsilon^2}{2\varepsilon} \log \frac{1+\varepsilon}{1-\varepsilon}\right],\ P = \frac{4\pi\,(2a^2 + b^2)}{3}$$

son las expresiones exacta y aproximada, respectivamente. Encontrar

($a$) el término inicial del error relativo,

($b$) el error relativo cuando $b = 0$.

¿Qué pasa con el signo del error relativo?

**8.** Comparar los ejemplos 6 y 7, y proponer una fórmula para el área de la superficie del elipsoide en general con semiejes $a$, $b$ y $c$.

¿Qué pasa con el signo del error?

**9.** [Sección 2.] A partir de la representación paramétrica de la elipse, $x = a \operatorname{sen} t$, $y = b$, cos $t$, mostrar que

$$E = 4a \int_0^{\pi/2} (1 - \varepsilon^2 \operatorname{sen}^2 t)^{1/2}\, dt$$

$$= 2\pi a \left[1 - \sum_1^{\infty} \left(\frac{1}{2}\ \frac{3}{4}\ \cdots\ \frac{2n-1}{2n}\right)^2 \frac{\varepsilon^{2n}}{2n-1}\right]$$

y derivar de aquí los términos dados sin prueba en la sección 2.

**10.** (Continúa.) Utilizando las extensiones con valores $\varepsilon$, probar que $E > P$ para $0 < \varepsilon \leqq 1$.

**11.** [Sección 2.] Determinar el número $\alpha$ de modo que la expresión

$$P'' = \alpha P + (1 - \alpha) P'$$

consiga la mejor aproximación posible a $E$ para $\varepsilon$ pequeña. (Es decir, que el orden del término inicial de $(P'' - E)/E$ sea lo más alto posible.)

**12.** (Continúa.) Investigar la aproximación para $P''$ siguiendo el método de la sección 2. ¡Inductivamente!

**13.** Dado un entero positivo $p$ y una secuencia de números positivos $a_1, a_2, a_3, \ldots, a_n, \ldots$ Mostrar que

$$\limsup_{n \to \infty} \left( \frac{a_1 + a_{n+p}}{a_n} \right)^n \geqq e^p.$$

**14.** (Continúa.) Indicar una secuencia $a_1, a_2, a_3, \ldots$, que consiga la igualdad.

**15.** *Explicar las regularidades observadas.* En física los descubrimientos surgen en dos etapas. Primero se observa cierta regularidad en los datos experimentales. Luego esta regularidad queda explicada como una consecuencia de alguna ley general. Cada una de estas etapas pueden ser recorridas por personas distintas, separadas por un largo intervalo de tiempo. Un magnífico ejemplo es el de Kepler y Newton: las regularidades en el movimiento de los planetas observadas por Kepler fueron explicadas por la ley de gravitación descubierta por Newton. Algo similar puede suceder en la investigación matemática, y aquí tenemos un claro ejemplo que requiere muy poco conocimiento preliminar.

La tabla normal de logaritmos vulgares de cuatro cifras registra 900 mantisas, las de los logaritmos de los enteros que van de 100 a 999. Sin haber atendido a la observación, podemos inclinarnos a pensar que las diez cifras, 0, 1, ..., 9 son igualmente frecuentes en estas tablas, pero la cosa no es así: ciertamente, no se suceden con la misma frecuencia que la *primera* cifra de la mantisa. Contando las mantisas que tienen el primer número igual, obtenemos la Tabla I. (Comprobémoslo.)

TABLA I

MANTISAS CON EL PRIMER NUMERO IGUAL EN LOGARITMOS DE CUATRO

| Primer número | Número de mantisas | Razones |
|---|---|---|
| 0 | 26 | |
| | | 1.269 |
| 1 | 33 | |
| | | 1.242 |
| 2 | 41 | |
| | | 1.268 |
| 3 | 52 | |
| | | 1.250 |
| 4 | 65 | |
| | | 1.262 |
| 5 | 82 | |
| | | 1.256 |
| 6 | 103 | |
| | | 1.252 |
| 7 | 129 | |
| | | 1.271 |
| 8 | 164 | |
| | | 1.250 |
| 9 | 205 | |
| | TOTAL. 900 | |

Mirando la segunda columna de la tabla podemos darnos cuenta de que dos números consecutivos cualesquiera tienen, aproximadamente, la misma razón. Esto nos induce a calcular estas razones con unos pocos decimales: pueden verse en la última columna de la tabla I.

¿Por qué son estas razones aproximadamente iguales? Intentemos percibir alguna regularidad precisa tras la aproximada que hemos observado. Los números de la segunda columna de la tabla I son aproximadamente los términos de una progresión geométrica. ¿Podemos descubrir una progresión geométrica exacta a la que puedan ser referidos los términos de la progresión aproximada? [La razón de la progresión exacta será, quizá, alguna clase de promedio entre las razones de la última columna de la tabla I.]

**16.** *Clasificar los hechos observados.* Una gran parte del trabajo del naturalista tiene como fin describir y clasificar los objetos que observa. Tal tipo de trabajo ha predominado mucho tiempo después de

cuando la principal actividad de los naturalistas consistía en describir nuevos géneros y especies de plantas y animales, y volver a clasificar las especies y géneros conocidos. No sólo las plantas y los animales son descritos y clasificados por los naturalistas, sino también otros objetos, sobre todo minerales; la clasificación de cristales está basada sobre su simetría. Una buena clasificación es importante; reduce toda la variedad observable a unos pocos tipos claramente caracterizados y ordenados.

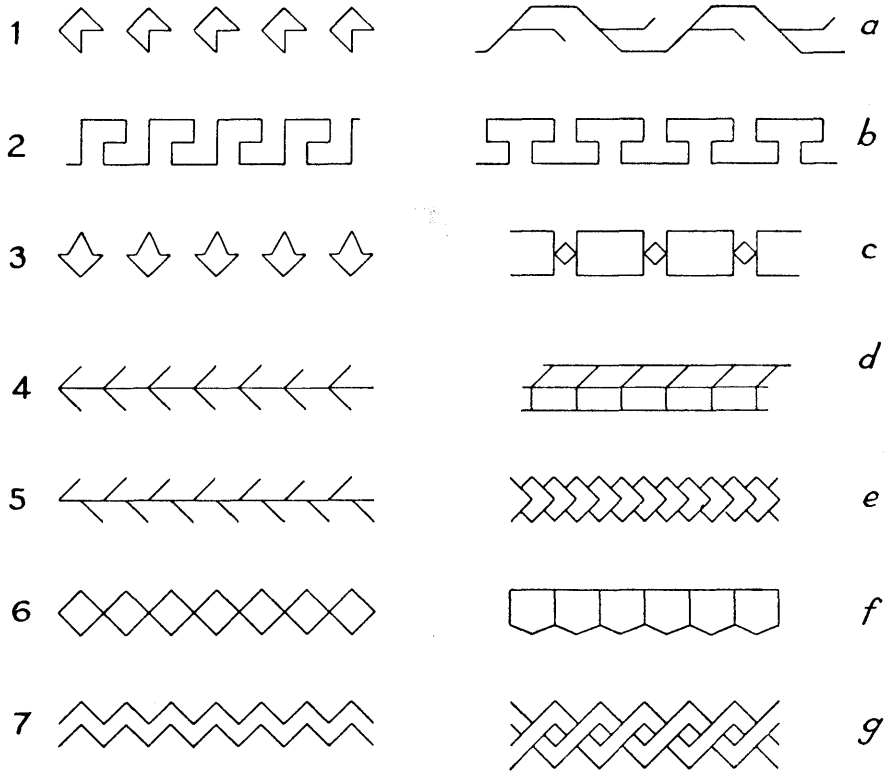



Fig. 5.2. Simetrías de frisos.

El matemático no tiene muy a menudo la oportunidad de entregarse a la descripción y clasificación, pero alguna vez sucede.

Si usted está familiarizado con algunas pocas y sencillas nociones de geometría plana (línea de simetría, centro de simetría) puede divertirse con ornamentos. La figura 5.2 exhibe catorce bandas ornamentales, cada una de las cuales es engendrada por una figura sencilla, periódicamente repetida a lo largo de una línea recta (horizontal). Llamemos a una banda tal «friso». Emparejemos cada friso del lado izquierdo de la figura 5.2 (marcado con un número) con uno del lado derecho (marcado

con una letra) de modo que los dos frisos emparejados tengan *el mismo tipo de simetría.* Además, examinar las bandas ornamentales que encontremos en toda clase de objetos, o en viejas obras arquitectónicas, e intentar emparejarlas con un friso de la figura 5.2. Finalmente, dar una lista completa de los varios tipos de simetría que un friso puede tener y hacer una descripción exhaustiva de cada uno de estos tipos. [Considerar un friso como infinitamente largo, en ambas direcciones, y la figura generante como periódicamente repetida una infinidad de veces. Observar que el término «tipo de simetría» no ha sido formalmente definido: llegar a una adecuada interpretación de este término es una parte importante de la tarea del estudiante.]

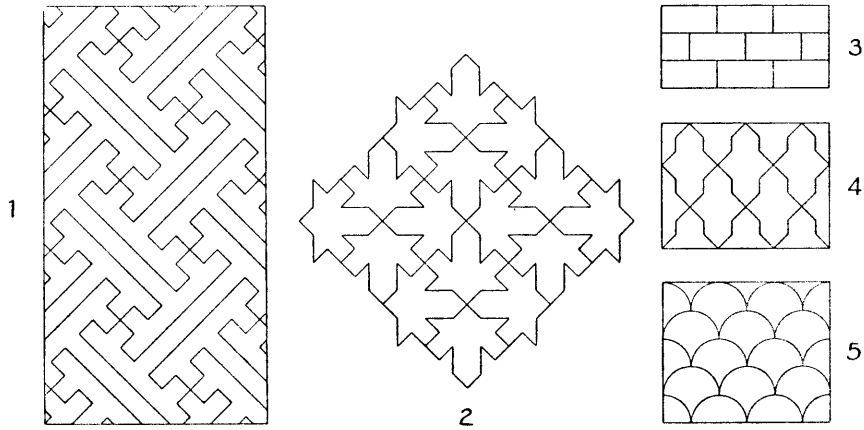



Fig. 5.3. Simetrías de papel para empapelar.

**17.** Encontrar en la figura 5.3 dos ornamentos que tengan el mismo tipo de simetría. Cada ornamento debe ser concebido como cubriendo todo el plano con sus modelos repetidos.

**18.** *¿Cuál es la diferencia?* Las veintiséis letras capitales han sido divididas en cinco tandas como sigue:

A M T U V W Y
B C D E K
N S Z
H I O X
F G J L P Q R.

¿Cuál es la diferencia? ¿Cuál podría ser una sencilla base para la clasificación exhibida? [Mirar las cinco ecuaciones:

$$y = x^2, \quad y^2 = x, \quad y = x^3, \quad x^2 + 2y^2 = 1, \quad y = x + x^4.$$

¿Cuál es la diferencia?]

# Un enunciado más general

> *[Euler] prefería instruir a sus alumnos con la pequeña satisfacción de sorprenderlos. El pensaba no haber hecho bastante por la ciencia si no hubiese añadido a los descubrimientos con que la enriqueció la íntegra exposición de las ideas que le llevaron a esos descubrimientos.*—CONDORCET.

## 1. EULER

De todos los matemáticos con que yo estoy de alguna manera familiarizado Euler me parece con mucho ser el más importante para nuestra investigación. Maestro de la investigación inductiva en matemáticas, hizo importantes descubrimientos (sobre series infinitas en la teoría de los números, y en otras ramas de las matemáticas) por inducción, es decir, por observación; intuiciones atrevidas y astutas verificaciones. En este aspecto, sin embargo, Euler no es único; otros matemáticos, de mucha y de poca importancia, utilizaron la inducción extensamente en su trabajo.

Euler me parece único en otro aspecto: él se tomó la molestia de presentar la evidencia inductiva cuidadosamente, en detalle y con buen orden; la presenta convincentemente, pero con honestidad, como lo haría un genuino científico. Su presentación es «la íntegra exposición de las ideas que le condujeron a esos descubrimientos», y tiene un encanto distintivo. Naturalmente que, como todos los autores, trata de impresionar a sus lectores; pero, como los buenos autores, sólo trata de impresionarlos con las cosas que antes le han impresionado a él mismo.

La próxima sección trae una muestra de los escritos de Euler. La memoria elegida puede ser leída con muy poco conocimiento previo y está enteramente dedicada a la exposición de un argumento inductivo.

## 2. LA MEMORIA DE EULER

La memoria de Euler que damos aquí *in extenso* no tiene más que algunas alteraciones secundarias que la harán más accesible al lector moderno *.

* El original está en francés; véase la *Opera Omnia*, de EULER, ser. 1, vol. 2, páginas 241-253. Las alteraciones consisten en una notación diferente (véase la nota 2),

*Descubrimiento de una extraordinaria ley de los números concerniente a la suma de sus divisores.*

1. Hasta ahora los matemáticos han intentado en vano descubrir algún orden en la secuencia de los números primos y tenemos toda la razón en creer que hay en ello algún misterio que la mente humana nunca logrará penetrar. Para convencernos basta con echar una ojeada a la tabla de los números primos, que algunos se han molestado en calcular hasta los 100.000, y percibir que no hay orden ni regla. Esto es mucho más sorprendente dado que la aritmética nos da reglas definidas con cuya ayuda podemos continuar la secuencia de los primos tan lejos como queramos, sin notar, sin embargo, la menor traza de orden. Yo mismo me encuentro muy lejos de esta meta, pero he descubierto una ley extremadamente extraña que gobierna las sumas de los divisores de los números enteros, que, a primera vista, aparece tan irregular como la secuencia de los primos y que, en cierto sentido, comprende también a la última. Esta ley, que explicaré en seguida, es, en mi opinión, tanto más digna de nota cuanto que podemos estar seguros de su verdad sin necesidad de dar una perfecta demostración de ella. Sin embargo, yo presentaré tal evidencia de modo que sea considerada como casi equivalente a una rigurosa demostración.

2. Un número primo no tiene más divisores que la unidad y sí mismo, y esto distingue al primo de cualquier otro número. Así, 7 es primo, porque es divisible por 1 y por sí mismo solamente. Cualquier otro número que tenga, además de la unidad y sí mismo, otros divisores recibe el nombre de compuesto, como, por ejemplo, 15, que tiene, además de 1 y 15, los divisores 3 y 5. Por tanto, generalmente, si el número $p$ es primo, sólo puede ser divisible por 1 y $p$; pero si $p$ fuese compuesto tendría, aparte 1 y $p$, otros divisores. Por tanto, en el primer caso la suma de sus divisores será $1 + p$, mientras en el último excederá a $p$. Como tengo que considerar la suma de los divisores de varios números, usaré * el signo $\sigma(n)$ para denotar la suma de los divisores del número $n$. Así, $\sigma(12)$ significa la suma de todos los divisores de 12, que son 1, 2, 3, 4, 6 y 12; por tanto, $\sigma(12) = 28$. Del mismo modo podemos ver que $\sigma(60) = 168$ y $\sigma(100) = 217$. Sin embargo, puesto que la unidad

---

en la ordenación de una tabla (explicada en la nota 3), en ligeros cambios que afectan algunas fórmulas y en quitar una repetición de argumentos en el último número 13 de la memoria. El lector interesado podrá consultar fácilmente el original.

* Euler fue el primero en introducir un símbolo para la suma de los divisores; él usó $\int n$, no el moderno $\sigma(n)$ del texto.

es solamente divisible por sí misma, $\sigma(1) = 1$. Ahora bien, 0 (cero) es divisible por todos los números. Por tanto, $\sigma$ (0) sería, en rigor, infinito. (Sin embargo, asignaré al último un valor finito, diferente en diferentes casos, y esto lo hará muy útil.)

3 Habiendo definido el significado del símbolo $\sigma(n)$, como arriba, vemos claramente que si $p$ es un primo, $\sigma(p) = 1 + p$. Pero $\sigma(1) = 1$ (y no $1 + 1$); de aquí que 1 sea excluido de la secuencia de los primos; 1 es el comienzo de los enteros, ni primo ni compuesto. Si, sin embargo, $n$ es compuesto, $\sigma(n)$ es más grande que $1 + n$.

En este caso podemos fácilmente encontrar $\sigma(n)$ a partir de los factores de $n$. Si $a, b, c, d, \ldots$ son primos diferentes, vemos fácilmente que

$$\begin{aligned}\sigma(ab) &= 1 + a + b + ab = (1 + a)(1 + b) = \sigma(a)\,\sigma(b),\\ \sigma(abc) &= (1 + a)(1 + b)(1 + c) = \sigma(a)\sigma(b)\sigma(c),\\ \sigma(abcd) &= \sigma(a)\sigma(b)\sigma(c)\sigma(d),\end{aligned}$$

y así sucesivamente. Necesitamos reglas particulares para los valores de los primos, como

$$\sigma(a^2) = 1 + a + a^2 = \frac{a^3 - 1}{a - 1}$$

$$\sigma(a^3) = 1 + a + a^2 + a^3 = \frac{a^4 - 1}{a - 1}$$

y, generalmente,

$$\sigma(a^n) = \frac{a^{n+1} - 1}{a - 1}.$$

Usando ésta podemos encontrar la suma de los divisores de cualquier número, compuesto de cualquier manera. Esto vemos con las fórmulas

$$\begin{aligned}\sigma(a^2b) &= \sigma(a^2)\sigma(b)\\ \sigma(a^3b^2) &= \sigma(a^3)\sigma(b^2)\\ \sigma(a^3b^4c) &= \sigma(a^3)\sigma(b^4)\sigma(c)\end{aligned}$$

y, generalmente,

$$\sigma(a^\alpha b^\beta c^\gamma d^\delta e^\varepsilon) = \sigma(a^\alpha)\sigma(b^\beta)\sigma(c^\gamma)\sigma(d^\delta)\sigma(e^\varepsilon).$$

Por ejemplo, para encontrar $\sigma(360)$ proponemos, ya que los factores de 360 son $2^3 \cdot 3^2 \cdot 5$,

$$\sigma(360) = \sigma(2^3)\sigma(3^2)\sigma(5) = 15 \cdot 13 \cdot 6 = 1170.$$

4. Con objeto de mostrar la secuencia de las sumas de los divisores he añadido por mi cuenta las siguiente tabla *, que contiene las sumas de los divisores de todos los enteros desde 1 a 99.

| *n* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | — | 1 | **3** | **4** | 7 | **6** | 12 | **8** | 15 | 13 |
| 10 | 18 | **12** | 28 | **14** | 24 | 24 | 31 | **18** | 39 | **20** |
| 20 | 42 | 32 | 36 | **24** | 60 | 31 | 42 | 40 | 56 | **30** |
| 30 | 72 | **32** | 63 | 48 | 54 | 48 | 91 | **38** | 60 | 56 |
| 40 | 90 | **42** | 96 | **44** | 84 | 78 | 72 | **48** | 124 | 57 |
| 50 | 93 | 72 | 98 | **54** | 120 | 72 | 120 | 80 | 90 | **60** |
| 60 | 168 | **62** | 96 | 104 | 127 | 84 | 144 | **68** | 126 | 96 |
| 70 | 144 | **72** | 195 | **74** | 114 | 124 | 140 | 96 | 168 | **80** |
| 80 | 186 | 121 | 126 | **84** | 224 | 108 | 132 | 120 | 180 | **90** |
| 90 | 234 | 112 | 168 | 128 | 144 | 120 | 252 | **98** | 171 | 156 |

Si examinamos un poco la secuencia de estos números estaremos a punto de desesperar. No parece posible descubrir el menor orden. La irregularidad de los primos está tan profundamente envuelta en ella que parece imposible encontrar la ley que gobierna esta secuencia, a menos que conozcamos la ley que gobierna la secuencia de los mismos primos. Parece así que la secuencia que tenemos ante nosotros es más misteriosa que la de los mismos primos.

5. Sin embargo, he observado que esta secuencia está sujeta a una ley completamente definida y se la puede considerar como una secuencia *recurrente*. Esta expresión matemática significa que cada término puede ser calculado por los términos precedentes, de acuerdo a una regla invariable. En efecto, si $\sigma(n)$ denota cualquier término de la secuencia, y $\sigma(n-1)$, $\sigma(n-2)$, $\sigma(n-3)$, $\sigma(n-4$, $\sigma(n-5)$, ..., los términos precedentes, digo que el valor $\sigma(n)$ puede siempre ser compuesto con algunos de los precedentes como prescrito por la fórmula siguiente:

$$\begin{aligned}\sigma(n) = {} & \sigma(n-1) + \sigma(n-2) - \sigma(n-5) - \sigma(n-7) \\ & + \sigma(n-12) + \sigma(n-15) - \sigma(n-22) - \sigma(n-26) \\ & + \sigma(n-35) + \sigma(n-40) - \sigma(n-51) - \sigma(n-57) \\ & + \sigma(n-70) + \sigma(n-77) - \sigma(n-92) - \sigma(n-100). \\ & + \ldots\end{aligned}$$

Sobre esta fórmula debemos hacer las siguientes observaciones.

* El número que se halla en la intersección de la fila 60 y la columna 7, es decir, 68, es $\sigma(67)$. Si $p$ es primo, $\sigma(p)$ está impreso en caracteres gruesos. Esta ordenación de la tabla es un poco más concisa que la original.

I. En la secuencia de los signos + y — cada uno comporta dos de sucesión.

II. La ley de los números 1, 2, 5, 7, 12, 15, ..., que hemos deducido del número propuesto $n$, llegará a ser más clara si observamos sus diferencias:

Números: 1, 2, 5, 7, 12, 15, 22, 26, 35, 40, 51, 57, 70, 77, 92, 100, ...
Diferencias: **1**, 3, **2**, 5, **3**, 7, **4**, 9, **5**, 11, **6**, 13, **7**, 15, **8**, ...

En efecto, aquí tenemos alternadamente todos los enteros 1, 2, 3, 4, 5, 6, ..., y los números impares 3, 5, 7, 9, 11, ..., y de aquí podemos continuar la secuencia de estos números tan lejos como queramos.

III. Aunque esta secuencia conduce al infinito, debemos tomar en cada caso solamente los términos para los que los números bajo el signo σ son todavía positivos y omitir la σ para valores negativos.

IV. Si el signo σ(0) reaparece en la fórmula, debemos, puesto que el valor es en sí mismo indeterminado, sustituir para σ(0) el número $n$ propuesto.

6. Tras estas observaciones no es difícil aplicar la fórmula a cualquier caso particular, y así cualquiera puede convencerse a sí mismo de su verdad con tantos ejemplos como desee desarrollar. Y puesto que debo admitir que no estoy en posición de dar una rigurosa demostración, lo justificaré con un gran número de ejemplos.

$$\begin{aligned}
\sigma(1) &= \sigma(0) &&= 1 &&= 1\\
\sigma(2) &= \sigma(1) + \sigma(0) &&= 1 + 2 &&= 3\\
\sigma(3) &= \sigma(2) + \sigma(1) &&= 3 + 1 &&= 4\\
\sigma(4) &= \sigma(3) + \sigma(2) &&= 4 + 3 &&= 7\\
\sigma(5) &= \sigma(4) + \sigma(3) - \sigma(0) &&= 7 + 4 - 5 &&= 6\\
\sigma(6) &= \sigma(5) + \sigma(4) - \sigma(1) &&= 6 + 7 - 1 &&= 12\\
\sigma(7) &= \sigma(6) + \sigma(5) - \sigma(2) - \sigma(0) &&= 12 + 6 - 3 - 7 &&= 8\\
\sigma(8) &= \sigma(7) + \sigma(6) - \sigma(3) - \sigma(1) &&= 8 + 12 - 4 - 1 &&= 15\\
\sigma(9) &= \sigma(8) + \sigma(7) - \sigma(4) - \sigma(2) &&= 15 + 8 - 7 - 3 &&= 13\\
\sigma(10) &= \sigma(9) + \sigma(8) - \sigma(5) - \sigma(3) &&= 13 + 15 - 6 - 4 &&= 18\\
\sigma(11) &= \sigma(10) + \sigma(9) - \sigma(6) - \sigma(4) &&= 18 + 13 - 12 - 7 &&= 12\\
\sigma(12) &= \sigma(11) + \sigma(10) - \sigma(7) - \sigma(5) + \sigma(0) &&= 12 + 18 - 8 - 6 + 12 &&= 28\\
\sigma(13) &= \sigma(12) + \sigma(11) - \sigma(8) - \sigma(6) + \sigma(1) &&= 28 + 12 - 15 - 12 + 1 &&= 14\\
\sigma(14) &= \sigma(13) + \sigma(12) - \sigma(9) - \sigma(7) + \sigma(2) &&= 14 + 28 - 13 - 8 + 3 &&= 24\\
\sigma(15) &= \sigma(14) + \sigma(13) - \sigma(10) - \sigma(8) + \sigma(3) + \sigma(0) &&= 24 + 14 - 18 - 15 + 4 + 15 &&= 24\\
\sigma(16) &= \sigma(15) + \sigma(14) - \sigma(11) - \sigma(9) + 0(4) + \sigma(1) &&= 24 + 24 - 12 - 13 + 7 + 1 &&= 31\\
\sigma(17) &= \sigma(16) + \sigma(15) - \sigma(12) - \sigma(10) + \sigma(5) + \sigma(2) &&= 31 + 24 - 28 - 18 + 6 + 3 &&= 18\\
\sigma(18) &= \sigma(17) + \sigma(16) - \sigma(13) - \sigma(11) + \sigma(6) + \sigma(3) &&= 18 + 31 - 14 - 12 + 12 + 4 &&= 39\\
\sigma(19) &= \sigma(18) + \sigma(17) - \sigma(14) - \sigma(12) + \sigma(7) + \sigma(4) &&= 39 + 18 - 24 - 28 + 8 + 7 &&= 20\\
\sigma(20) &= \sigma(19) + \sigma(18) - \sigma(15) - \sigma(13) + \sigma(8) + \sigma(5) &&= 20 + 39 - 24 - 14 + 15 + 6 &&= 42
\end{aligned}$$

Creo que estos ejemplos son suficientes para desanimar a cualquiera que se atreva a pensar que el acuerdo de mi regla con la realidad es por mera casualidad.

7. Sin embargo, aún alguien podría dudar si la ley de los números 1, 2, 5, 7, 12, 15, ..., que hemos deducido es precisamente esa que hemos indicado, puesto que los ejemplos dados implican sólo los primeros seis de esos números. Así, la ley podría aparecer aún como insuficientemente establecida y, por tanto, todavía daré algunos ejemplos con números mayores.

I. Dado el número 101, encontrar la suma de sus divisores. Tenemos

$$\begin{aligned}
\sigma(101) &= \sigma(100) + \sigma(99) - \sigma(96) - \sigma(94) \\
&+ \sigma(89) + \sigma(86) - \sigma(79) - \sigma(75) \\
&+ \sigma(66) + \sigma(61) - \sigma(50) - \sigma(44) \\
&+ \sigma(31) + \sigma(24) - \sigma(9) - \sigma(1) \\
&= 217 + 156 - 252 - 144 \\
&+ 90 + 132 - 80 - 124 \\
&+ 144 + 62 - 93 - 84 \\
&+ 32 + 60 - 13 - 1 \\
&= 893 - 791 \\
&= 102
\end{aligned}$$

y de aquí podríamos concluir, si no nos fuese anteriormente conocido, que 101 es un número primo.

II. Dado el número 301, encontrar la suma de sus divisores. Tenemos

$$\begin{aligned}
&\text{dif.} \qquad 1 \qquad 3 \qquad 2 \qquad 5 \\
\sigma(301) &= \sigma(300) + \sigma(299) - \sigma(296) - \sigma(294) + \\
&\qquad 3 \qquad 7 \qquad 4 \qquad 9 \\
&+ \sigma(289) + \sigma(286) - \sigma(279) - \sigma(275) + \\
&\qquad 5 \qquad 11 \qquad 6 \qquad 13 \\
&+ \sigma(266) + \sigma(261) - \sigma(250) - \sigma(244) + \\
&\qquad 7 \qquad 15 \qquad 8 \qquad 17 \\
&+ \sigma(231) + \sigma(224) - \sigma(209) - \sigma(201) + \\
&\qquad 9 \qquad 19 \qquad 10 \qquad 21 \\
&+ \sigma(184) + \sigma(175) - \sigma(156) - \sigma(146) + \\
&\qquad 11 \qquad 23 \qquad 12 \qquad 25 \\
&+ \sigma(125) + \sigma(114) - \sigma(91) - \sigma(79) + \\
&\qquad 13 \qquad 27 \qquad 14 \\
&+ \sigma(54) + \sigma(41) - \sigma(14) - \sigma(0).
\end{aligned}$$

Vemos por este ejemplo cómo podemos, usando las diferencias, continuar la fórmula tan lejos como sea necesario en cada caso. Realizando los cálculos encontramos

$$\sigma(301) = 4939 - 4587 = 352.$$

De aquí concluimos que 301 no es primo. En efecto, $301 = 7 \cdot 43$ y obtenemos

$$\sigma(301) = \sigma(7)\sigma(43) = 8 \cdot 44 = 352$$

como la regla ha mostrado.

8. Los ejemplos que hemos desarrollado eliminan, indudablemente, cualquier escrúpulo que hubiéramos tenido sobre la verdad de mi fórmula. Ahora, esta bella propiedad de los números es tanto más sorprendente cuanto que no percibimos ninguna conexión inteligible entre la estructura de mi fórmula y la naturaleza de los divisores cuya suma nos concierne. La secuencia de los números 1, 2, 5, 7, 12, 15, ..., no parece tener ninguna relación con la cuestión que discutimos. Además, como la ley de estos números está «interrumpida» y, de hecho, son la mezcla de dos secuencias con una ley regular, de 1, 5, 12, 22, 35, 51, ... y 2, 7, 15, 26, 40, 57, ..., no esperaríamos que tal irregularidad apareciese en el análisis. La falta de demostración puede aumentar la sorpresa todavía más, puesto que parece totalmente imposible lograr descubrir tal propiedad sin ser guiado por algún método fiable que tomara el lugar de una prueba perfecta. Por mi parte, confieso que no acerté con este descubrimiento por casualidad, sino que otra proposición me abrió el camino a esta bella propiedad —otra proposición de la misma naturaleza, que puede ser aceptada como cierta, aunque me encuentre incapaz de probarla—. Y a pesar de que aquí consideramos la naturaleza de los números a los que el cálculo infinitesimal no parece aplicarse, yo he alcanzado mi conclusión por diferenciaciones y otros mecanismos. Me gustaría que alguien encontrase un camino más corto y natural, en el que la consideración del camino que yo he seguido pueda, quizá, servir de ayuda.

9. Considerando las divisiones de los números, examiné, hace tiempo, la expresión

$$(1-x)(1-x^2)(1-x^3)(1-x^4)(1-x^5)(1-x^6)(1-x^7)(1-x^8)\ldots,$$

en que el producto se asume que es infinito. Con objeto de ver qué clase de series resultarán multipliqué, en realidad, un gran número de factores y encontré

$$1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - x^{35} - x^{40} + \ldots$$

Los exponentes de $x$ son los mismos que entran en la fórmula de arriba; también aquí los signos + y — se suceden de dos en dos. Basta emprender esta multiplicación y continuarla tan lejos como estimemos necesario para convencernos de la verdad de esta serie. Sin embargo, no hay otra evidencia de ella que una larga inducción realizada tan lejos que no podemos dudar sobre la ley que gobierna la formación de estos términos y sus exponentes. Por mi parte, he buscado en vano una rigurosa demostración de la ecuación entre la serie y el producto infinito de arriba $(1 - x)(1 - x^2)(1 - x^3)\ldots$, y he propuesto la misma cuestión a algunos de mis amigos cuya habilidad en estas materias conozco muy bien, pero todos se han mostrado de acuerdo conmigo sobre la verdad de esta transformación del producto en una serie, sin ser capaces de descubrir ninguna posibilidad de demostración. Nos encontramos, pues, ante una verdad conocida, aunque no demostrada, que podemos escribir así:

$$s = (1 - x)(1 - x^2)(1 - x^3)(1 - x^4)(1 - x^5)(1 - x^6)\ldots$$

cuya cantidad $s$ puede expresarse como sigue:

$$s = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - x^{35} - x^{40} + \ldots$$

Cada uno de nosotros puede convencerse a sí mismo de esta verdad realizando la multiplicación tan lejos como desee; y nos ha de parecer imposible que la ley que hemos visto cumplirse en veinte términos, por ejemplo, deje de cumplirse en los términos siguientes.

10. Así como hemos descubierto que estas dos expresiones infinitas son iguales, aunque nos haya sido imposible demostrar su verdad, todas las conclusiones que sean deducidas de ellas serán de la misma naturaleza, es decir, ciertas aunque no demostradas. Ahora bien, si una de estas conclusiones fuese demostrada equivaldría a obtener una clave para la demostración de esa ecuación; y con este exclusivo propósito he manejado esas dos expresiones en muy diferentes maneras, que me condujeron entre otros descubrimientos a ese que he explicado arriba; su verdad, por tanto, debe ser tan cierta como la de la ecuación entre dos expresiones infinitas. Yo procedí como sigue. Siendo las dos expresiones dadas

I. $s = (1 - x)(1 - x^2)(1 - x^3)(1 - x^4)(1 - x^5)(1 - x^6)(1 - x^7)\ldots$

II. $s = 1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - x^{35} - x^{40} + \ldots$

iguales, me desembarecé de los factores en la primera sacando logaritmos

$$\log s = \log(1-x) + \log(1-x^2) + \log(1-x^3) + \log(1-x^4) + \ldots$$

Con el fin de desembarazarse de los logaritmos diferencié y obtuve la ecuación

$$\frac{1}{s}\frac{ds}{dx} = -\frac{1}{1-x} - \frac{2x}{1-x^2} - \frac{3x^2}{1-x^3} - \frac{4x^3}{1-x^4} - \frac{5x^4}{1-x^5} - \ldots$$

o

$$-\frac{x}{s}\frac{ds}{dx} = \frac{x}{1-x} + \frac{2x^2}{1-x^2} + \frac{3x^3}{1-x^3} + \frac{4x^4}{1-x^4} + \frac{5x^5}{1-x^5} + \ldots$$

De la segunda expresión para $s$, como serie infinita, obtenemos otro valor para la misma cantidad

$$-\frac{x}{s}\frac{ds}{dx} = \frac{x+2x^2-5x^5-7x^7+12x^{12}+15x^{15}-22x^{22}-26x^{26}+\ldots}{1-x-x^2+x^5+x^7-x^{12}-x^{15}+x^{22}+x^{26}-\ldots}.$$

11. Pongamos

$$-\frac{x}{s}\frac{ds}{dx} = t.$$

Tenemos arriba dos expresiones para la cantidad $t$. En la primera expresión desarrollo cada término en una serie geométrica y obtengo

$$\begin{array}{rcccccccccc}
t = x + & x^2 + & x^3 + & x^4 + & x^5 + & x^6 + & x^7 + & x^8 + \ldots \\
& + 2x^2 & & + 2x^4 & & + 2x^6 & & + 2x^8 + \ldots \\
& & + 3x^3 & & & + 3x^6 & & + \ldots \\
& & & + 4x^4 & & & & + 4x^8 + \ldots \\
& & & & + 5x^5 & & & + \ldots \\
& & & & & + 6x^6 & & + \ldots \\
& & & & & & + 7x^7 & + \ldots \\
& & & & & & & + 8x^8 + \ldots
\end{array}$$

Aquí vemos fácilmente que cada valor de $x$ surge tantas veces como su exponente tiene divisores, y que cada divisor aparece como un coeficiente del mismo valor de $x$. Por tanto, si recogemos los términos

con valores iguales, el coeficiente de cada valor de $x$ será la suma de los divisores de su exponente. Y, en consecuencia, utilizando la notación de arriba $\sigma(n)$ para la suma de los divisores de $n$, obtenemos

$$t = \sigma(1)x + \sigma(2)x^2 + \sigma(3)x^3 + \sigma(4)x^4 + \sigma(5)x^5 + \ldots$$

La ley de la serie es manifiesta. Y, aunque pudiera parecer que la inducción estaba envuelta en la determinación de los coeficientes, podemos fácilmente convencernos de que esta ley es una consecuencia necesaria.

12. En virtud de la definición de $t$, la última fórmula del número 10 puede ser escrita como sigue:

$$t(1 - x - x^2 + x^5 + x^7 - x^{12} - x^{15} + x^{22} + x^{26} - \ldots)$$
$$- x - 2x^2 + 5x^5 + 7x^7 - 12x^{12} - 15x^{15} + 22x^{22} + 26x^{26} - \ldots = 0.$$

Sustituyendo $t$ por el valor obtenido al final del número 11, encontramos

$$\begin{aligned} 0 = {} & \sigma(1)x + \sigma(2)x^2 + \sigma(3)x^3 + \sigma(4)x^4 + \sigma(5)x^5 + \sigma(6)x^6 + \ldots \\ & - x - \sigma(1)x^2 - \sigma(2)x^3 - \sigma(3)x^4 - \sigma(4)x^5 - \sigma(5)x^6 - \ldots \\ & - 2x^2 - \sigma(1)x^3 - \sigma(2)x^4 - \sigma(3)x^5 - \sigma(4)x^6 - \ldots \\ & + 5x^5 + \sigma(1)x^6 + \ldots \end{aligned}$$

Recogiendo los términos encontraremos el coeficiente para cualquier valor dado de $x$. Este coeficiente se compone de varios términos. Primero viene la suma de los divisores del exponente de $x$ y luego las sumas de los divisores de algunos números precedentes, obtenidos de ese exponente por sustracción sucesiva de 1, 2, 5, 7, 12, 15, 22, 26, ... Finalmente, si pertenece a esta secuencia, el mismo exponente surge. No necesitamos explicar otra vez los signos de los términos enumerados. Por tanto, en general, el coeficiente de $x^n$ es

$$\sigma(n) - \sigma(n-1) - \sigma(n-2) + \sigma(n-5) + \sigma(n-7) - \sigma(n-12) - \sigma(n-15) + \ldots$$

Esta serie puede prolongarse todo lo que se quiera mientras los números bajo el signo $\sigma$ no sean negativos. Sin embargo, si surge el término $\sigma(0)$, debemos sustituirlo por $n$.

13. Puesto que la suma de la serie infinita considerada en el parágrafo número 12 es 0, cualquiera que sea el valor de $x$, el coeficiente de cada valor único de $x$ ha de ser necesariamente 0. De aquí obtenemos la ley que expliqué más arriba en el número 5; me refiero a la ley que

gobierna la suma de los divisores y nos capacita para calcular todos los números. En el desarrollo precedente podemos percibir alguna razón para los signos, alguna para la secuencia de los números

$$1, 2, 5, 7, 12, 15, 22, 26, 35, 40, 51, 57, 70, 77, \ldots$$

y, especialmente, una razón para sustituir por $n$ el número $\sigma(0)$, que parece ser el más extraño rasgo de mi regla. Este razonamiento, aunque todavía muy lejos de una demostración perfecta, dejará pocas dudas sobre la muy extraordinaria ley que he explicado aquí.

## 3. Transición a un punto de vista más general

El anterior texto de Euler es extraordinariamente instructivo. Podemos aprender mucho de él, tanto en matemáticas como en psicología de la invención y en razonamiento inductivo. Los ejemplos y comentarios al fin de este capítulo nos darán una oportunidad de examinar algunas ideas matemáticas de Euler, pero ahora deseamos concentrarnos sobre su argumento inductivo.

El teorema investigado por Euler es digno de nota en varios aspectos y es de gran interés matemático incluso hoy en día. Sin embargo, no nos interesa aquí tanto el contenido matemático del teorema como las razones que indujeron a Euler a creer en el teorema cuando aún no había sido probado. Con objeto de comprender mejor la naturaleza de estas razones ignoraré el contenido matemático de la memoria de Euler y daré un esquemático bosquejo de ella, enfatizando cierto aspecto general de su argumento inductivo.

Como desatenderemos el contenido matemático de los varios teoremas discutidos, será mejor designarlos con letras, como $T$, $T^*$, $C_1$, $C_2$, ..., $C^*$, $C_2^*$, ... El lector puede ignorar el significado de estas letras completamente. Sin embargo, en el caso de querer reconocerlas en el texto de Euler, he aquí la clave.

$T$ es el teorema

$$(1-x)(1-x^2)(1-x^3)\ldots = 1-x-x^2+x^5+x^7-x^{12}-x^{15}+\ldots$$

La ley de los números 1, 2, 5, 7, 12, 15, ... es explicado en la sección 2, número 5, II.

$C_n$ es la afirmación de que el coeficiente de $x^n$ es el mismo a ambos lados de la ecuación anterior. Por ejemplo, $C_6$ afirma que extendiendo el producto del lado izquierdo encontraremos que el coeficiente de $x^6$ es 0. Observar que $C_n$ es una consecuencia del teorema $T$.

$C_n^*$ es la ecuación

$$\sigma(n) = \sigma(n-1) + \sigma(n-2) - \sigma(n-5) - \sigma(n-7) + \ldots$$

explicada extensamente en la sección 2, número 5. Por ejemplo, $C_6^*$ afirma que

$$\sigma(6) = \sigma(5) + \sigma(4) - \sigma(1).$$

$T^*$ es la «muy extraordinaria ley», afirmativa de que $C_1^*$, $C_2^*$, $C_3^*$, ... son todos ciertos. Observar que $C_n^*$ es una consecuencia, un caso particular de $T^*$.

## 4. Bosquejo esquemático de la memoria de Euler *

El *teorema T* es de tal naturaleza que podemos estar seguros de su verdad sin dar una perfecta demostración de él. Sin embargo, presentaremos una tal evidencia de él que puede ser mirado como una casi rigurosa demostración.

*El teorema T incluye un número infinito de casos particulares:* $C_1$, $C_2$, $C_3$, ... *A la inversa, el conjunto infinito de estos casos particulares* $C_1$, $C_2$, $C_3$, ... *es equivalente al teorema T. Podemos hallar por simple cálculo si* $C_1$ *es cierto o no. Otro cálculo simple determina si* $C_2$ *es cierto o no, y de modo semejante para* $C_3$, *y así sucesivamente. Yo he hecho este cálculo y he encontrado que* $C_1$, $C_2$, $C_3$, ..., $C_{40}$ *son todos ciertos.* Basta emprender *estos cálculos* y continuarlos hasta donde consideremos conveniente para convencernos de la verdad de esta secuencia *indefinidamente continuada.* Carezco de otra evidencia que no sea la inducción realizada hasta un punto en que no tengo ninguna duda que la ley de $C_1$, $C_2$, ... *son casos particulares.* En vano hemos buscado una demostración rigurosa del *teorema T*, y la misma cuestión presentada a amigos cuya habilidad en estas materias conozco me ha mostrado de acuerdo con ellos sobre la verdad del teorema, aunque también sobre la incapacidad de demostrarlo. Así, conocemos una verdad, aunque no haya sido todavía demostrada; cada uno de nosotros podemos convencernos igualmente por el *cálculo real de los casos* $C_1$, $C_2$, $C_3$, ..., tantos como deseemos; parece imposible que la verdad descubierta para veinte términos, por ejemplo, no lo sea también en los casos no observados.

Así como hemos descubierto la verdad del *teorema T*, aunque nos haya sido imposible demostrarla, de la misma manera todas las conclu-

* Este bosquejo fue por primera vez publicado en mi artículo «El razonamiento heurístico y la teoría de la probabilidad», *Am. Math. Monthly*, vol. 48, 1941, págs. 450-465. Las palabras en cursiva indican frases que no son de Euler.

siones que sean deducidas de él tendrán la misma naturaleza; serán ciertas aunque indemostradas. Ahora bien, si una de estas conclusiones llega a ser demostrada, tenemos con ello una clave para demostrar recíprocamente el *teorema T*; con este propósito en la mente hemos maniobrado el *teorema T* de muchas maneras, hasta descubrir, entre otros, el *teorema T** cuya verdad debe ser tan cierta como la del *teorema T*.

*Los teoremas T y T* son equivalentes; ambos son verdaderos o falsos; o se mantienen o caen juntos. Como T, el teorema T* incluye un número infinito de casos particulares* $C_1^*$, $C_2^*$, $C_3^*$, ..., *y esta secuencia de casos particulares es equivalente al teorema T*. Aquí de nuevo un simple cálculo nos muestra si* $C_1^*$ *es cierto o no. De modo semejante es posible demostrar si* $C_2^*$ *es cierto o no, y así sucesivamente.* No es difícil aplicar el teorema *T** a cualquier caso particular dado, y así cualquiera puede satisfacerse a sí mismo con tantos ejemplos como desee desarrollar. Y puesto que debo admitir que no estoy en posición de dar una rigurosa demostración, lo justificaré con un número suficientemente grande de ejemplos, $C_1^*$, $C_2^*$, $C_3^*$, ... $C_{20}^*$. Creo que estos ejemplos son suficientes para desanimar a cualquiera de pensar que es pura casualidad que mi regla esté de acuerdo con la verdad.

Si alguien duda todavía que la ley es esa que he indicado, aún puedo darle muchos más ejemplos con números más grandes. *Por cálculo, he encontrado que* $C_{101}^*$, *y* $C_{301}^*$, *son verdaderos, de modo que el teorema T* es válido también para estos casos tan alejados de los que examiné al principio.* Estos ejemplos que he desarrollado despejarán, indudablemente, todos los escrúpulos que pudiesen quedar sobre la verdad de los *teoremas T y T**.

## EJEMPLOS Y COMENTARIOS

Al descubrir su «muy extraordinaria ley de los números» Euler «llegó a su conclusión por diferenciaciones y otros mecanismos», aunque «el cálculo infinitesimal no parece aplicarse a la naturaleza de los enteros». Con el fin de comprender el método de Euler le aplicaremos a ejemplos similares. Empecemos por dar nombre a su principal «mecanismo» o instrumento matemático.

**1.** *Funciones generantes.* El resultado del número 11 de la memoria de Euler en notación moderna es éste:

$$\sum_{n=1}^{\infty} \frac{nx^n}{1-x^n} = \sigma(1)x + \sigma(2)x^2 + \ldots + \sigma(n)x^n + \ldots$$

La derecha es una serie de valores de $x$. El coeficiente de $x^n$ en esta serie es $\sigma(n)$, la suma de los divisores de $n$. Ambos lados de la ecuación representan la misma función de $x$. La extensión de esta función en valores de $x$ «genera» la secuencia $\sigma(1)$, $\sigma(2)$, ..., $\sigma(n)$, ... y, por ello, la llamamos *función generante* de $\sigma(n)$. En general, si

$$f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n + \ldots,$$

decimos que $f(x)$ es la función generante de $a_n$, o la función que genera la secuencia $a_0, a_1, a_2, \ldots, a_n, \ldots$

El nombre de «función generante» se debe a Laplace. Sin embargo, sin darle un nombre, el mismo Euler usó el mecanismo de las funciones generantes mucho tiempo antes de Laplace, en varias memorias, de las que hemos visto un ejemplo en la sección 2. Aplicó este instrumento matemático a varios problemas de análisis combinatorio y de teoría de los números.

Una función generante es un mecanismo de algún modo parecido a un saco. En vez de transportar muchos pequeños objetos separadamente, lo que podría ser embarazoso, los metemos en un saco a todos juntos y así sólo tenemos que transportar un objeto: el saco. De modo semejante, en lugar de habérnoslas con cada término de la secuencia $a_0, a_1, a_2, \ldots, a_n, \ldots$, individualmente, los metemos a todos en una serie de valor $\Sigma a_nx^n$, de modo que tengamos un solo objeto matemático que manejar.

**2.** hállese la función generante de $n$. O, lo que es lo mismo, hállese la suma de la serie $\Sigma nx^n$.

**3.** Siendo dado que $f(x)$ genera la secuencia $a_0, a_1, a_2, \ldots, a_n, \ldots$ hállese la función generante de la secuencia

$$0a_0, 1a_1, 2a_2, \ldots, na_n, \ldots$$

**4.** Siendo dado que $f(x)$ genera la secuencia $a_0, a_1, a_2, \ldots, a_n, \ldots$, hállese la función generante de la secuencia

$$0, a_0, a_1, \ldots, a_{n-1}, \ldots$$

**5.** Siendo dado que $f(x)$ es la función generante de $a_n$, hállese la función generante de

$$s_n = a_0 + a_1 + a_2 + \ldots + a_n.$$

**6.** Siendo dado que $f(x)$ y $g(x)$ son las funciones generantes de $a_n$ y $b_n$, respectivamente, hállese la función generante de

$$c_n = a_0b_n + a_1b_{n-1} + a_2b_{n-2} + \ldots + a_nb_0.$$

**7.** ***Un problema combinatorio en geometría plana.*** Un polígono convexo con $n$ lados es dividido en $n-2$ triángulos por $n-3$ diagonales; ver la figura 6.1. Llamemos $D_n$ al número de divisiones.

Encontrar $D_n$ para $n = 3, 4, 5, 6$.

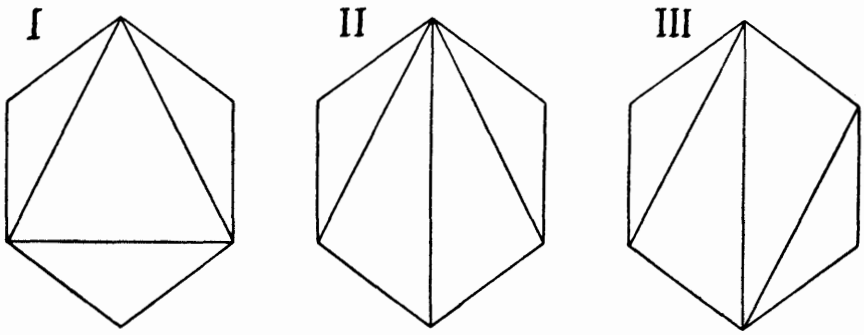



Fig. 6.1. Tres tipos de división de un hexágono.

**8.** (Continúa.) No es fácil intuir una expresión general y explícita para $D_n$ sobre la base de los valores numéricos considerados en el ejemplo 7. Sin embargo, la secuencia $D_3$, $D_4$, $D_5$, ... es una sentencia «recurrente» en el siguiente y muy general sentido: cada uno de los términos puede ser calculado a partir de los términos precedentes de acuerdo con una regla invariable, una «fórmula de recursión». (Véase la memoria de Euler, número 5.)

Definir $D_2 = 1$, y mostrar que para $n \geq 3$

$$D_n = D_2 D_{n-1} + D_3 D_{n-2} + D_4 D_{n-3} + \ldots + D_{n-1} D_2.$$

[Comprobar los primeros casos. Referirse a la figura 6.2.]

**9.** (Continúa.) La derivación de una expresión explícita para $D_n$ a partir de la fórmula de recursión del ejemplo 8 no es obvia. Sin embargo, considerar la función generante

$$g(x) = D_2 x^2 + D_3 x^3 + D_4 x^4 + \ldots + D_n x^n + \ldots$$

Mostrar que $g(x)$ satisface una ecuación cuadrática y derivar de ahí que para $n = 3, 4, 5, 6, \ldots$

$$D_n = \frac{2}{2}\,\frac{6}{3}\,\frac{10}{4}\,\frac{14}{5}\,\ldots\,\frac{4n-10}{n-1}.$$

**10.** *Sumas de cuadrados.* Recuerden la definición de $R_k(n)$ (ejemplo 4.1), desarróllenla a $n = 0$ en el caso $R_k(0) = 1$ (un desarrollo razonable), introduzcan la función generante

$$\sum_{n=0}^{\infty} R_k(n)x^n = R_k(0) + R_k(1)x + R_k(2)x^2 + \ldots,$$

y muestren que

$$\sum_{n=0}^{\infty} R_3(n)x^n = (1 + 2x + 2x^4 + 2x^9 + \ldots)^3.$$

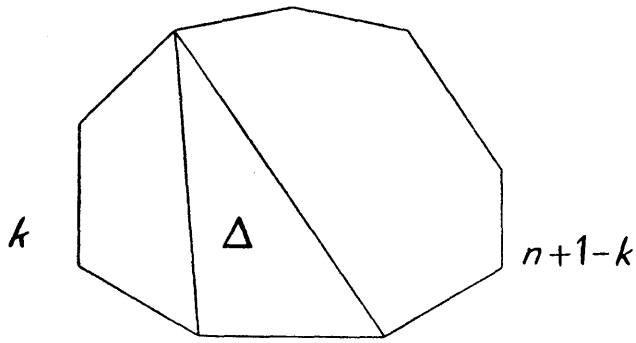



Fig. 6.2. Iniciación de la división de un polígono con $n$ lados.

[¿Qué es $R_3(n)$? El número de soluciones de la ecuación

$$u^2 + v^2 + w^2 = n$$

en enteros $u$, $v$ y $w$, positivos, negativos, o 0.

¿Cuál puede ser el papel de la serie del lado derecho de la ecuación que se le ha propuesto probar?

$$1 + 2x + 2x^4 + 2x^9 + \ldots = \sum_{u=-\infty}^{\infty} x^{u^2} = \sum_{n=0}^{\infty} R_1(n)x^n.$$

¿Cómo concebir la ecuación del lado derecho a la que se dirige? Quizá así:

$$\Sigma x^{u^2} \cdot \Sigma x^{u^2} \cdot \Sigma x^{u^2}.]$$

**11.** Generalicen el resultado del ejemplo 10.

**12.** Recuerden la definición de $S_k(n)$ (ejemplo 4.1) y expresen la función generante

$$\sum_{n=1}^{\infty} S_k(n)x^n.$$

**13.** Usen el ejemplo 11 para probar que para $n \geq 1$, $R_2(n)$ es divisible por 4, $R_4(n)$ por 8 y $R_8(n)$ por 16. (El resultado fue ya utilizado en el capítulo IV, Tablas II y III.)

**14.** Utilicen el ejemplo 12 para probar que

$$S_2(n) = 0 \text{ si } n \text{ no es de la forma } 8m + 2,$$
$$S_4(n) = 0 \text{ si } n \text{ no es de la forma } 8m + 4,$$
$$S_8(n) = 0 \text{ si } n \text{ no es de la forma } 8m.$$

**15.** Utilicen el ejemplo 11 para probar que

$$R_{k+l}(n) = R_k(0)R_l(n) + R_k(1)R_l(n-1) + \ldots + R_k(n)R_l(0).$$

**16.** Prueben que

$$S_{k+l}(n) = S_k(1)S_l(n-1) + S_k(2)S_l(n-2) + \ldots + S_k(n-1)S_l(1).$$

**17.** Propongan un método sencillo para calcular la Tabla III del capítulo IV a partir de las tablas I y II del mismo capítulo.

**18.** Establezcamos $\sigma_k(n)$ para la suma de los valores $k$ de los divisores de $n$. Por ejemplo,

$$\sigma_3(15) = 1^3 + 3^3 + 5^3 + 15^3 = 3528;$$

$\sigma_1(n) = \sigma(n)$.

(1) Muestren que las conjeturas encontradas en la sección 4.6 y el ejemplo 4.23 implican

$$\sigma(1)\sigma(2u-1) + \sigma(3)\sigma(2u-3) + \ldots + \sigma(2u-1)\sigma(1) = \sigma_3(u),$$

donde $u$ denota un entero impar.

(2) Ensayen casos particulares de la relación encontrada numéricamente en (1).

(3) ¿Cómo una verificación tal influye en la confianza sobre las conjeturas a partir de las cuales se ha derivado la relación verificada?

**19.** *Otra fórmula periódica.* Consideremos las funciones generantes

$$G = \sum_{m=1}^{\infty} S_1(m)x^m, \qquad H = \sum_{m=1}^{\infty} S_4(m)x^m.$$

Pongamos

$$S_4(4u) = s_u,$$

donde $u$ es un entero impar. Luego,

$$G = x + x^9 + x^{25} + \ldots + x^{(2n-1)^4} + \ldots ,$$
$$H = s_1x^4 + s_3x^{12} + s_5x^{20} + \ldots + s_{2n-1}x^{8n-4} + \ldots,$$
$$G^4 = H$$

por los ejemplos 14 y 12. De la última cuestión derivamos, sacando logaritmos y diferenciando, que

$$4 \log G = \log H,$$
$$\frac{4G'}{G} = \frac{H'}{H'}$$
$$G \,.\, xH' = 4 \,.\, xG' \,.\, H,$$

$$\begin{aligned}(x + \ x^9 + \ \ x^{25} + \ldots) \ \ &(4s_1x^4 + 12s_3x^{12} + 20s_5x^{20} + \ldots)\\ = 4(x + 9x^9 + 25x^{25} + \ldots) \ \ &(\ s_1x^4 + \ \ s_3x^{12} + \ \ s_5x^{20} + \ldots).\end{aligned}$$

Comparando los coeficientes de $x^5$, $x^{13}$, $x^{21}$, . . . a ambos lados de la ecuación anterior, encontramos, tras algunas elementales operaciones, las siguientes relaciones:

$$\begin{aligned}
0s_1 &= 0\\
1s_3 - 4s_1 &= 0\\
2s_5 - 3s_3 &= 0\\
3s_7 - 2s_5 - 12s_1 &= 0\\
4s_9 - 1s_7 - 11s_3 &= 0\\
5s_{11} - 10s_5 &= 0\\
6s_{13} + 1s_{11} - 9s_7 - 24s_1 &= 0\\
7s_{15} + 2s_{13} - 8s_9 - 23s_3 &= 0\\
8s_{17} + 3s_{15} - 7s_{11} - 22s_5 &= 0\\
9s_{19} + 4s_{17} - 6s_{13} - 21s_7 &= 0\\
10s_{21} + 5s_{19} - 5s_{15} - 20s_9 - 40s_1 &= 0\\
11s_{23} + 6s_{21} - 4s_{17} - 19s_{11} - 39s_3 &= 0\\
&\ldots\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

La primera ecuación de este sistema es vacua y ha sido expuesta aquí sólo para enfatizar la ley general. Sin embargo, sabemos que $s_1 = 1$. Sabiendo esto, obtenemos la próxima ecuación $s_3$, y con éste la ecuación siguiente $s_5$. Y así sucesivamente, podemos calcular el sistema de términos de la secuencia $s_1, s_3, s_5, \ldots$ tan lejos como queramos, una tras otra, «recurrentemente» o periódicamente.

El sistema sigue una estructura observable. Hay 1 ecuación que contiene 1 de las cantidades $s_1$, $s_3$, $s_5$, ..., 2 que contienen 2 de ellas, 3 que contienen 3, etc., etc. Los coeficientes de cada columna aumentan en 1, mientras que los suscritos aumentan en 2 al pasar de una fila a otra. El suscrito a la cabeza de cada columna es 1 y el coeficiente es —4 multiplicado por el primer coeficiente de la misma fila.

Podemos resumir el sistema total en una ecuación (fórmula de recursión); escribanla.

**20.** *Otra ley muy extraordinaria de los números referente a la suma de sus divisores.* Si la conjetura de la sección 4.6 establece

$$s_{2n-1} + S_4(4(2n-1)) = \sigma(2n-1)$$

y el ejemplo 19 proporciona una fórmula de recursión conectando los términos de la secuencia $\sigma(1)$, $\sigma(3)$, $\sigma(5)$, $\sigma(7)$, ... que es, en muchos aspectos, asombrosamente parecida a la fórmula de Euler.

Desarrollen en detalle y verifiquen numéricamente los primeros casos de la indicada fórmula de recursión.

**21.** Para nosotros hay también una semejanza heurística entre la fórmula de recursión de Euler para $\sigma(n)$ (sec. 2) y la anterior fórmula de recursión para $\sigma(2n-1)$ (ejemplo 20). Para nosotros esta última es una conjetura. Hemos derivado esta conjetura, como Euler derivó la suya, «por diferenciación y otros mecanismos» a partir de otra conjetura.

Muestren que la fórmula de recursión para $\sigma(2n-1)$ indicada por el ejemplo 20 es equivalente a la ecuación

$$S_4(4(2n-1)) = \sigma(2n-1),$$

a la que hemos llegado en la sección 4.6. Es decir, si una de las dos aserciones es cierta, la otra es necesariamente cierta.

**22.** Generalicen el ejemplo 19.

**23.** Ingenien un método para calcular $R_8(n)$ independientemente de $R_4(n)$.

**24.** *Cómo Euler erró un descubrimiento.* El método ilustrado por los ejemplos 19 y 23, y establecido en general en el ejemplo 22, se debe a Euler *. Al inventar este método Euler apuntaba al problema de las cuatro raíces y algunos problemas relacionados. De hecho, él aplicó su método al problema de las cuatro raíces e investigó inductivamente el número de representaciones, pero fracasó en descubrir la ley

* *Opera Omnia*, ser. 1, vol. 4, págs. 125-135.

que gobierna $R_4(n)$, la cual no es, después de todo, tan difícil de descubrir inductivamente (ejemplos 4.10-4.15). ¿Cómo sucedió?

Al examinar la ecuación

$$n = x^2 + y^2 + z^2 + w^2$$

podemos elegir varios puntos de partida, especialmente los siguientes:

(1) Admitir para $x$, $y$, $z$ y $w$ sólo enteros no-negativos.

(2) Admitir para $x$, $y$, $z$ y $w$ todos enteros, positivos, negativos y cero.

El segundo punto puede ser menos obvio, pero conduce a $R_4(n)$ y a la conexión entre $R_4(n)$ y los divisores de $n$. El primer punto es más obvio, pero el número de soluciones no parece tener ninguna sencilla propiedad digna de nota. Euler eligió el punto de vista (1), no el (2), y explicó su método explicado en el ejemplo 22 a

$$(1 + x + x^4 + x^9 + \ldots)^4,$$

no a

$$(1 + 2x + 2x^4 + 2x^9 + \ldots)^4,$$

y así fue desaprovechado un gran descubrimiento. Es muy instructivo comparar estas dos líneas de investigación que parecen iguales al principio, pero de las cuales una es maravillosamente fructífera, mientras la otra resulta casi completamente estéril.

Las propiedades $R_4(n)$, $S_4(n)$, $R_8(n)$ y $S_8(n)$ investigadas en el capítulo IV (ejemplos 4.10-4.15, secciones 4.3-4.6, ejemplos 4.18-4.23) fueron descubiertas por Jacobi, no inductivamente, sino como corolarios incidentales de sus investigaciones sobre las funciones elípticas. Varias pruebas de estos teoremas se han encontrado desde entonces, pero ninguna lo bastante elemental ni recta *.

**25.** *Una generalización del teorema de Euler sobre* $\sigma(n)$. Dado $k$, tenemos

$$\prod_{n=1}^{\infty} (1 - x^n)^k = 1 - \sum_{n=1}^{\infty} a_n x^n$$

y muestren que, para $n = 1, 2, 3, \ldots$

$$\sigma(n) = \sum_{m=1}^{n-1} a_m \sigma(n - m) + n a_n / k.$$

¿Qué caso particular nos proporciona el teorema de Euler de la sección 2?

---

* Véase también, para más referencia, G. H. HARDY y E. M. WRIGHT, *An introduction to the theory of numbers*, Oxford, 948, cap. XX.

# La inducción matemática

*El método de Jacques Bernouille es importante también para el naturalista. Encontramos lo que parece ser una propiedad* A *del concepto* B *por observación de los casos* $C_1$, $C_2$, $C_3$, ... *Aprendemos del método de Bernouille que no debíamos atribuir tal propiedad* A, *encontrada por inducción incompleta, no matemática, al concepto* B, *a menos que* percibamos *que* A *está* ligado *a las características de* B *y es* independiente *de la variación de los casos. Como en muchos otros puntos, la matemática ofrece aquí un modelo a la ciencia natural.*—ERNST MACH *.

## 1. La fase inductiva

De nuevo empecemos con un ejemplo.

Hay muy poca dificultad en encontrar la suma de los primeros $n$ enteros. Tomemos aquí por garantía la fórmula

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2},$$

que puede ser descubierta y probada de muchas maneras **. Es muy pesado encontrar una fórmula para la suma de los primeros $n$ cuadrados

$$1 + 4 + 9 + 16 + \ldots + n^2.$$

No hay dificultad en calcular esta suma para pequeños valores de $n$, pero no es tan fácil desentrañar una regla. Es bastante natural, sin embargo, buscar alguna especie de paralelismo entre las dos sumas y observarlas juntas:

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | ... |
|---|---|---|---|---|---|---|---|
| $1 + 2 + \ldots + n$ | 1 | 3 | 6 | 10 | 15 | 21 | ... |
| $1^2 + 2^2 + \ldots + n^2$ | 1 | 5 | 14 | 30 | 55 | 91 | ... |

* *Erkenntnis und Irrtum*, 4.ª ed., 1920, pág. 312.

** Ver *How to Solve It*, pág. 107.

¿Cómo están relacionadas las dos últimas filas? Podemos averiguarlo examinando su razón:

$$\begin{array}{cc} n & 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad \ldots \\ \dfrac{1^2 + 2^2 + \ldots + n^2}{1 + 2 + \ldots + n} & 1 \quad \dfrac{5}{3} \quad \dfrac{7}{3} \quad 3 \quad \dfrac{11}{3} \quad \dfrac{13}{3} \quad \ldots \end{array}$$

Aquí la regla es obvia y es casi imposible no percibirla si la escribimos como sigue:

$$\frac{3}{3} \quad \frac{5}{3} \quad \frac{7}{3} \quad \frac{9}{3} \quad \frac{11}{3} \quad \frac{13}{3}.$$

No podemos reprimirnos de formular la conjetura que

$$\frac{1^2 + 2^2 + \ldots + n^2}{1 + 2 + \ldots + n} = \frac{2n + 1}{3}.$$

Utilizando el valor del denominador del lado izquierdo nos vemos conducidos a establecer la conjetura de esta manera:

$$1^2 + 2^2 + 3^2 + \ldots + n^2 = \frac{n(n + 1)\,(2n + 1)}{6}.$$

¿Es esto cierto? Es decir, ¿es cierto en general? La fórmula es realmente verdadera en los casos particulares $n = 1, 2, 3, 4, 5, 6$, que la sugieren. ¿Es también verdadero el caso próximo $n = 7$? La conjetura nos conduce a predecir que

$$1 + 4 + 9 + 16 + 25 + 36 + 49 = \frac{7 \cdot 8 \cdot 15}{6},$$

y, en efecto, ambos lados son iguales a 140.

Podríamos, por supuesto, seguir con el próximo caso $n = 8$ y ensayarlo, pero la tentación no es demasiado fuerte. Estamos inclinados a creer de cualquier modo que la fórmula se verificará en el próximo caso, y por ello esta verificación añadirá poco a nuestra confianza —tan poco que no merece la pena—. ¿Cómo podríamos ensayar la conjetura de modo más eficiente?

Si la conjetura es absolutamente cierta, será independiente de la variación de los casos y se mantendrá válida en la transición de un caso a otro.

Supuestamente,

$$1 + 4 + \ldots + n^2 = \frac{n(n + 1)\,(2n + 1)}{6}.$$

Sin embargo, si esta fórmula es cierta en general, *valdrá también para el próximo caso;* tendremos

$$1 + 4 + \ldots + n^2 + (n + 1)^2 = \frac{(n + 1)\,(n + 2)\,(2n + 3)}{6}.$$

Aquí hay una oportunidad de comprobar eficientemente la conjetura: sustrayendo la línea superior de la más baja obtenemos

$$(n + 1)^2 = \frac{(n + 1)\,(n + 2)\,(2n + 3)}{6} - \frac{n(n + 1)\,(2n + 1)}{6}.$$

¿Es ésta consecuencia de la conjetura cierta?
Una fácil reordenación del lado derecho nos proporciona

$$\frac{n + 1}{6}\,[(n + 2)\,(2n + 3) - n(2n + 1)]$$

$$= \frac{n + 1}{6}\,[2n^2 + 3n + 4n + 6 - 2n^2 - n]$$

$$= \frac{n + 1}{6}\,[6n + 6]$$

$$= (n + 1)^2.$$

La consecuencia examinada es incontestablemente cierta, la conjetura ha pasado una prueba severa.

## 2. La fase demostrativa

La verificación de cualquier consecuencia aumenta nuestra confianza en la conjetura, pero la verificación de la consecuencia justamente examinada puede hacer más: *demostrar* la conjetura. Sólo necesitamos

un ligero cambio de nuestro punto de vista y volver a barajar nuestras observaciones.

Es *supuestamente* cierto que

$$1^2 + 2^2 + \ldots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Es *incontestablemente* cierto que

$$(n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6} - \frac{n(n+1)(2n+1)}{6}.$$

Es *consecuentemente* cierto que

$$1^2 + 2^2 + \ldots + n^2 + (n+1)^2 = \frac{(n+1)(n+2)(2n+3)}{6}$$

(hemos añadido las dos ecuaciones anteriores). Esto significa: Si nuestra conjetura es cierta para un cierto entero $n$, necesariamente ha de continuar siendo verdad para el próximo entero $n+1$.

Sabemos que la conjetura es cierta para $n = 1, 2, 3, 4, 5, 6, 7$. Siendo cierta para 7 debe serlo para el próximo entero 8; también debe ser cierta, si lo es para 8, para 9, para 10, para 11, y así sucesivamente. La conjetura es, pues, cierta para todos los enteros; hemos conseguido probarla en toda su generalidad.

### 3. Examinando transiciones

El último razonamiento de la sección anterior puede ser un poco simplificado. Es suficiente saber dos cosas sobre la conjetura:

Es cierta para $n = 1$.

Siendo cierta para $n$, lo es igualmente para $n+1$.

Luego, la conjetura es cierta para todos los enteros: cierta para 1 por tanto, también para 2; cierta para 2, por tanto, también para 3 y así sucesivamente.

Tenemos aquí un importante procedimiento de demostración. Podríamos llamarlo «pasando de $n$ a $n+1$», pero es usualmente conocido como «inducción matemática». Esta última designación es, sin embargo, muy inadecuada para un procedimiento de demostración, pues la inducción (en el significado que con más frecuencia se utiliza) proporciona solamente una inferencia plausible, no demostrativa.

¿Tiene la inducción matemática algo que ver con la inducción? Sí, y por ello, y no por su nombre, la consideramos aquí.

En nuestro anterior ejemplo el razonamiento demostrativo de la sección 2 completa el razonamiento inductivo de la sección 1, y esto es típico. La demostración de la sección 2 aparece como «un complemento matemático a la inducción», y si tomamos la «inducción matemática» como una abreviación de dicho sentido, el término resulta bastante adecuado después de todo. (Por tanto, tomémoslo en este sentido —no busquemos disputas con términos técnicos establecidos—.) La inducción matemática a menudo aparece como el último grado, o la fase final, de una investigación inductiva, y esta última fase con frecuencia utiliza sugerencias que habían aparecido en fases precedentes.

Otra razón aún mejor para considerar que la inducción matemática en el presente contexto está oculta en el pasaje de Ernst Mach citado al principio de este capítulo *. Examinando una conjetura investigamos los varios casos a que la conjetura puede ser aplicada. Nuestro objetivo es ver si la relación afirmada por la conjetura es *estable*, es decir, independiente de la variación de los casos e inamovible ante ellos. Nuestra atención se dirige, naturalmente, a la *transición* de un caso a otro. «Que por medio de fuerzas centrípetas los planetas quedan retenidos en ciertas órbitas, puede ser fácilmente comprendido si consideramos los movimientos de los proyectiles», dice Newton, y entonces él imagina una piedra que es proyectada cada vez con mayor velocidad inicial hasta que su trayectoria da la vuelta a la Tierra como la trayectoria de la Luna; véase el ejemplo 2.18 (4). Así, Newton visualiza una transición continua del movimiento de un proyectil al movimiento de un planeta. El considera la transición entre dos casos a los cuales la ley de gravitación universal, que intentaba probar, es igualmente aplicable. Cualquier principiante, que use la inducción matemática para probar algún teorema elemental, actúa igual que Newton en este aspecto: él considera la transición de $n$ a $n + 1$, la transición entre dos casos a los que el teorema que trata de probar es igualmente aplicable.

## 4. La técnica de la inducción matemática

Para ser un buen matemático, o un buen jugador, o bueno en lo que sea, debemos ser buenos intuidores. Al objeto de ser un buen intuidor, usted debería empezar por ser, naturalmente, listo, me parece. Aunque

* Mach creía que Jacques Bernouille había inventado el método de la inducción matemática, pero hay muchos motivos dignos de crédito para suponer que esta invención se debe a Pascal. Cf. Freudenthal, *Archives internationales d'Histoire des Sciences*, núm. 22, 1953, pág. 1737. También, cf. *Jacobi Bernouille Bassileensis Opera*, Ginebra, 1744, vol. I, págs. 282-283.

ser listo no es bastante. Usted debería examinar sus intuiciones, compararlas con los hechos, modificarlas si fuese necesario, y así adquirir una extensa (e intensa) experiencia de las intuiciones que fracasan y las que llegan a ser ciertas. Con tal experiencia como base será mucho más capaz de juzgar competentemente qué intuiciones tienen la oportunidad de ser correctas y cuáles no.

La inducción matemática es un procedimiento útil a menudo para verificar conjeturas matemáticas, a las que hemos llegado por algún procedimiento inductivo. Por tanto, si deseamos adquirir alguna experiencia en investigación matemática inductiva una cierta familiaridad con la técnica de la inducción matemática es deseable.

La presente sección y los siguientes ejemplos y comentarios darán alguna ayuda para adquirir esta técnica.

(1) *La fase inductiva.* Empezamos con un ejemplo muy parecido al discutido en las secciones 1 y 2. Deseamos expresar en alguna forma más corta otra suma conectada con los primeros $n$ cuadrados:

$$\frac{1}{3} + \frac{1}{15} + \frac{1}{35} + \cdots + \frac{1}{4n^2 - 1} =$$

$$\frac{1}{4 \cdot 1^2 - 1} + \frac{1}{4 \cdot 2^2 - 1} + \frac{1}{4 \cdot 3^2 - 1} + \cdots + \frac{1}{4n^2 - 1}$$

Calculemos esta suma en los primeros casos y tabulemos los resultados:

$$n = 1, \quad 2, \quad 3, \quad 4, \ldots$$

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} = \frac{1}{3}, \ \frac{2}{5}, \ \frac{3}{7}, \ \frac{4}{9}, \ \cdots$$

Hay una intuición obvia:

$$\frac{1}{3} + \frac{1}{15} + \frac{1}{35} + \cdots + \frac{1}{4n^2 - 1} = \frac{n}{2n + 1}.$$

Aprovechando nuestra experiencia de nuestro primer problema similar, ensayemos la conjetura correcta tan lejos como eficientemente podamos:

ensayemos la transición de $n$ a $n + 1$. Si nuestra conjetura es, en general, cierta, debe ser cierta para $n$ y para $n + 1$:

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} = \frac{n}{2n + 1},$$

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} + \frac{1}{4(n + 1)^2 - 1} + \frac{n + 1}{2n + 3}.$$

Por sustracción obtenemos

$$\frac{1}{4(n + 1)^2 - 1} = \frac{n + 1}{2n + 3} - \frac{n}{2n + 1}.$$

¿Es cierta esta consecuencia de nuestra conjetura? Transformemos ambos lados, acercando uno al otro:

$$\frac{1}{(2n + 2)^2 - 1} = \frac{2n^2 + 3n + 1 - 2n^2 - 3n}{(2n + 3)(2n + 1)}.$$

No es necesario saber mucha álgebra para ver que los dos lados de la última ecuación son, en realidad, idénticos. La consecuencia examinada es incontestablemente cierta.

(2) *La fase demostrativa.* Ahora barajemos de nuevo nuestras observaciones, como en nuestro anterior ejemplo, sección 2.

Supuestamente,

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} = \frac{n}{2n + 1}.$$

Incontestablemente,

$$\frac{1}{4(n + 1)^2 - 1} = \frac{n + 1}{2n + 3} - \frac{n}{2n + 1}.$$

Consecuentemente,

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} + \frac{1}{4(n + 1)^2 - 1} = \frac{n + 1}{2n + 3}.$$

La conjetura, supuesta ser cierta para $n$, será también cierta, como consecuencia de esta suposición, también para $n + 1$. Como es cierta para $n = 1$, es cierta en general.

(3) *Más breve.* Podemos aún gastar menos tiempo en la fase inductiva de nuestra solución. Habiendo concebido la conjetura, podríamos sospechar que la inducción matemática es adecuada para probarla. Luego, sin ningún ensayo, podíamos haber intentado aplicar la inducción matemática directamente, como sigue.

Supuestamente,

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2 - 1} = \frac{n}{2n + 1}.$$

Consecuentemente,

$$\frac{1}{3} + \frac{1}{15} + \cdots + \frac{1}{4n^2-1} + \frac{1}{(4n+1)^2-1} = \frac{n}{2n+1} +$$

$$+ \frac{1}{4(n+1)^2-1}$$

$$= \frac{n}{2n+1} + \frac{1}{(2n+2)^2-1}$$

$$= \frac{n(2n+3)+1}{(2n+1)(2n+3)}$$

$$= \frac{2n^2+3n+1}{(2n+1)(2n+3)}$$

$$= \frac{(2n+1)(n+1)}{(2n+1)(2n+3)}$$

$$= \frac{n+1}{2n=3}$$

y así logramos derivar para $n = 1$ la relación que habíamos supuesto para $n$. Esto es exactamente lo que se nos pedía hacer, y con ello hemos probado la conjetura.

Esta variante de la solución es menos redundante, pero quizá también menos natural que la primera, presentada bajo (1) y (2).

(4) *Aún más breve.* Podemos ver la solución casi de una ojeada si notamos que

$$\frac{1}{4n^2-1} = \frac{1}{(2n-1)(2n+1)} = \frac{1}{2}\left(\frac{1}{2n-1} - \frac{1}{2n+1}\right).$$

(Somos llevados a esta fórmula de modo muy natural, si estamos familiarizados con la descomposición de funciones racionales en fracciones parciales.) Poniendo $n = 1, 2, 3, \ldots n$ y sumando, tenemos

$$\frac{1}{4-1} + \frac{1}{16-1} + \frac{1}{36-1} + \cdots + \frac{1}{4n^2-1}$$

$$= \frac{1}{1\cdot 3} + \frac{1}{3\cdot 5} + \frac{1}{5\cdot 7} + \cdots + \frac{1}{(2n-1)(2n+1)}$$

$$= \frac{1}{2}\left[\left(\frac{1}{1} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) + \left(\frac{1}{5} - \frac{1}{7}\right) + \cdots + \right.$$

$$\left. + \left(\frac{1}{2n-1} - \frac{1}{2n+1}\right)\right]$$

$$= \frac{1}{2}\left[1 - \frac{1}{2n+1}\right]$$

$$= \frac{n}{2n+1}.$$

Lo que ha sucedido no es infrecuente. Un teorema probado por inducción matemática puede ser probado a menudo más brevemente por algún otro método. Del mismo modo, el examen cuidadoso de la prueba por inducción matemática puede llevar a una tal abreviación.

(5) *Otro ejemplo.* Examinemos dos números, sujetos a las desigualdades

$$0 < a < 1, \quad 0 < b < 1.$$

Luego, obviamente,

$$(1-a)(1-b) = 1 - a - b + ab > 1 - a - b.$$

Una generalización natural nos conduce a sospechar el siguiente enunciado: *Si* $n \geq 2$ *y* $0 < a_1 < 1, 0 < a_2 < 1, \ldots 0 < a_n < 1$, *entonces*

$$(1-a_1)(1-a_2)\ldots(1-a_n) > 1 - a_1 - a_2 - \ldots - a_n.$$

Utilicemos la inducción matemática para probar esto. Hemos visto que la desigualdad es cierta en el primer caso, que se asegura aplicar para $n = 2$. Por tanto, suponiendo que sea cierto para $n$, donde $n \geqq 2$, tenemos que derivar para $n + 1$.

Supuestamente,

$$(1 - a_1) \ldots (1 - a_n) > 1 - a_1 - \ldots - a_n$$

y sabemos que

$$0 < a_{n+1} < 1.$$

Consecuentemente,

$$\begin{aligned}(1 - a_1) \ldots (1 - a_n)(1 - a_{n+1}) &> (1 - a_1 - \ldots - a_n)(1 - a_{n+1}) \\ &= 1 - a_1 - \ldots - a_n - a_{n+1} + (a_1 + \ldots + a_n)\, a_{n+1} \\ &> 1 - a_1 - \ldots - a_n - a_{n+1}.\end{aligned}$$

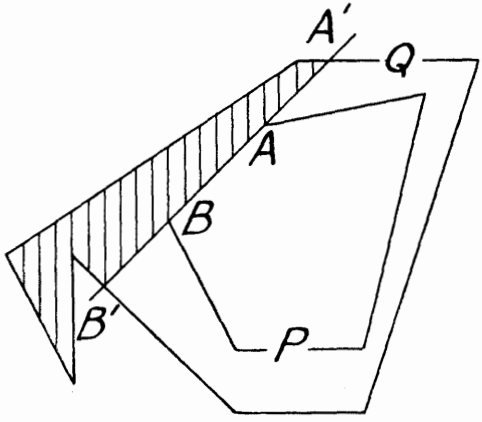



Fig. 7. 1. De $n$ a $n + 1$.

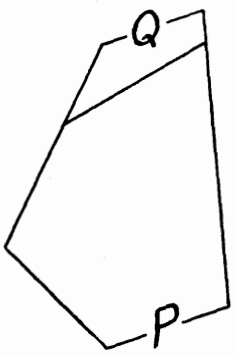



Fig. 7. 2 El El caso $n = 1$.

Notemos que la inducción matemática puede ser utilizada para probar proposiciones que aplicamos, no a todos los enteros positivos en absoluto, sino a todos los enteros positivos desde un cierto entero en adelante. Por ejemplo, el teorema probado concierne sólo a valores $n \geqq 2$.

(6) *¿Qué es n?* Ahora discutimos un teorema de geometría plana.

*Si el polígono P es convexo y está contenido en el polígono Q, el perímetro de P es más corto que el perímetro de Q.*

Que el área del polígono interno $P$ es menor que el área del polígono exterior $Q$ es obvio. Sin embargo, el teorema establecido no es obvio; sin la restricción de que $P$ es convexo sería falso.

La figura 7.1 muestra la idea esencial de la prueba. Separemos la parte sombreada del polígono $Q$; queda entonces un polígono $Q'$, una parte de $Q$, que tiene dos propiedades:

Primera, $Q'$ todavía contiene el polígono convexo $P$, que, siendo convexo, permanece a un lado de la línea recta $A'B'$, en el que queda el lado $AB$ de $P$.

Segundo, el perímetro de $Q'$ es más corto que el de $Q$. De hecho, el perímetro de $Q'$ difiere del de $Q$, en tanto que el primero contiene el segmento de línea recta que junta los puntos $A'$ y $B'$, y el último contiene, en cambio, una línea quebrada que junta los mismos puntos. Aun así, la línea recta es la distancia más corta entre los puntos $A'$ y $B'$.

Como hemos pasado de $Q$ a $Q'$, así podemos pasar de $Q'$ a otro polígono $Q''$. Obtendremos entonces una secuencia de polígonos $Q$, $Q'$, $Q''$, ... Cada polígono es incluido en el anterior y tiene un perímetro más corto que éste, y el último polígono en esta secuencia es $P$. Por tanto, el perímetro de $P$ es más corto que el de $Q$.

Reconozcamos la naturaleza de la prueba anterior: es, en efecto, una prueba de inducción matemática. Pero, ¿qué es $n$? ¿Con arreglo a qué cantidad es realizada la inducción?

Esta cuestión es seria. La inducción matemática es utilizada en varios dominios y algunas veces en muy difíciles e intrincadas cuestiones. Intentando encontrar una prueba oculta podemos enfrentarnos con una decisión crucial: ¿Qué será $n$? ¿Con arreglo a qué cantidad intentaremos la inducción matemática?

En la prueba anterior es aconsejable elegir como $n$ el *número de lados del polígono convexo interno que no pertenecen en ningún modo al perímetro del polígono exterior.* La figura 7.2 ilustra el caso $n = 1$. Dejamos al lector la tarea de averiguar qué es aconsejable llamar $n$ en la figura 7.1.

## EJEMPLOS Y COMENTARIOS

**1.** Observar que

$$\begin{aligned} 1 &= 1 \\ 1 - 4 &= -(1 + 2) \\ 1 - 4 + 9 &= 1 + 2 + 3 \\ 1 - 4 + 9 - 16 &= -(1 + 2 + 3 + 4). \end{aligned}$$

Intuir la ley general sugerida por estos ejemplos, expresarla en una conveniente notación matemática y probarla.

**2.** Probar las fórmulas explícitas para $P_n$, $S_n$ y $S_n^\infty$ intuidas en los ejemplos 3.13, 3.14 y 3.20, respectivamente. (Ejemplos 3.11, 3.12.)

**3.** Intuir una expresión para

$$1^3 + 2^3 + 3^3 + \ldots + n^3$$

y probarlas por inducción matemática. [Ejemplo 1.4.]

**4.** Intuir una expresión para

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\left(1 - \frac{1}{16}\right)\ldots\left(1 - \frac{1}{n^2}\right)$$

válida para $n \geqq 2$ y probarla por inducción matemática.

**5.** Intuir una expresión para

$$\left(1 - \frac{4}{1}\right)\left(1 - \frac{4}{9}\right)\left(1 - \frac{4}{25}\right)\ldots\left(1 - \frac{4}{(2n-1)^2}\right)$$

válida para $n \geqq 1$ y probarla por inducción matemática.

**6.** Generalizar la relación

$$\frac{x}{1+x} + \frac{2x^2}{1+x^2} + \frac{4x^4}{1+x^4} + \frac{8x^8}{1+x^8} = \frac{x}{1-x} - \frac{16x^{16}}{1-x^{16}}$$

y probar su generalización por inducción matemática.

**7.** Consideremos la operación que consiste en pasar de la secuencia

$$a_1, a_2, a_3, \ldots, a_n, \ldots$$

a la secuencia

$$s_1, s_2, s_3, \ldots, s_n, \ldots$$

con el término general

$$s_n = a_1 + a_2 + a_3 + \ldots + a_n.$$

Llamaremos esta operación (la formación de la secuencia $s_1, s_2, s_3, \ldots$) «suma de la secuencia $a_1, a_2, a_3, \ldots$». Con esta terminología expresamos un hecho ya observado (en el ejemplo 1.3) como sigue.

Podemos pasar de la secuencia de todos los positivos enteros 1, 2, 3, 4, ... a la secuencia de los cuadrados 1, 4, 9, 16, ... en dos pasos: (1) dejar fuera cada segundo término, (2) sumar la secuencia restante. En efecto, veamos la tabla:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | ... |
| 1 | | 3 | | 5 | | 7 | | 9 | | 11 | | 13 | ... |
| 1 | | 4 | | 9 | | 16 | | 25 | | 26 | | 49 | ... |

Pruébese esta aserción por inducción matemática.

**8.** (Continúa.) Usted puede pasar de la secuencia de todos los positivos enteros 1, 2, 3, 4, ... a la secuencia de los cubos 1, 8, 27, 64, ... en cuatro pasos: (1) dejar fuera cada tercer término, (2) sumar la secuencia restante, (3) dejar fuera cada segundo término, (4) sumar la secuencia restante. Pruebe esto por inducción matemática después de haber examinado la tabla:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | ... |
| 1 | 2 | | 4 | 5 | | 7 | 8 | | 10 | 11 | | 13 | ... |
| 1 | 3 | | 7 | 12 | | 19 | 27 | | 37 | 48 | | 61 | ... |
| 1 | | | 7 | | | 19 | | | 37 | | | 61 | ... |
| 1 | | | 8 | | | 27 | | | 64 | | | 125 | ... |

**9.** (Continúa.) Usted puede pasar de la secuencia de todos los positivos enteros 1, 2, 3, 4, ... a la secuencia de los cuatro valores 1, 16, 81, 256, ... en seis pasos, visibles en la tabla:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | ... |
| 1 | 2 | 3 | | 5 | 6 | 7 | | 9 | 10 | 11 | | 13 | ... |
| 1 | 3 | 6 | | 11 | 17 | 24 | | 33 | 43 | 54 | | 67 | ... |
| 1 | 3 | | | 11 | 17 | | | 33 | 43 | | | 67 | ... |
| 1 | 4 | | | 15 | 32 | | | 65 | 108 | | | 175 | ... |
| 1 | | | | 15 | | | | 65 | | | | 175 | ... |
| 1 | | | | 16 | | | | 81 | | | | 256 | ... |

¿Qué sugieren estos hechos?

**10.** Observar que

$$\begin{aligned}
1 &= 1\\
1 - 5 &= -4\\
1 - 5 + 10 &= 6\\
1 - 5 + 10 - 10 &= -4\\
1 - 5 + 10 - 10 + 5 &= 1\\
1 - 5 + 10 - 10 + 5 - 1 &= 0
\end{aligned}$$

de donde somos llevados al enunciado general

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \cdots + (-1)^k \binom{n}{k} = (-1)^k \binom{n-1}{k}$$

para $0 < k < n$, $n = 1, 2, 3, \ldots$

Al probar esto por inducción matemática, ¿usted prefiere proceder de $n$ a $n + 1$ o más bien de $k$ a $k + 1$?

**11.** En un campeonato de tenis hay $2n$ participantes. En la primera vuelta del campeonato cada participante juega sólo una vez; así, hay $n$ partidas, cada una de las cuales supone un par de jugadores. Mostrar que el emparejamiento de la primera vuelta puede ser ordenado exactamente en

$$1 \cdot 3 \cdot 5 \cdot 7 \cdots (2n - 1)$$

maneras diferentes.

**12.** *Probar más puede traer menos problemas.* Tomemos

$$\frac{1}{1 - x} = f_0(x)$$

y definamos la secuencia $f_0(x), f_1(x), f_2(x), \ldots$ con la condición de que

$$f_{n+1}(x) = x \frac{df_n(x)}{dx}$$

para $n = 0, 1, 2, 3, \ldots$ (Tal definición es llamada *recursiva:* para encontrar $f_{n+1}$ tenemos que volver a $f_n$.) Observando

$$f_1(x) = \frac{x}{(1 - x)^2}, \quad f_2(x) = \frac{x + x^2}{(1 - x)^3}, \quad f_3(x) = \frac{x + 4x^2 + x^3}{(1 - x)^4},$$

probar por inducción matemática que, para $n \geq 1$, el numerador de $f_n(x)$ es un polinomio, cuyo término constante es 0 y enteros positivos los otros coeficientes.

**13.** (Continúa.) Encontrar por inducción y probar por inducción matemática las restantes propiedades de $f_n(x)$.

**14.** *Equilibre su teorema.* La proposición típica $A$, accesible de probar por inducción matemática, tiene una infinidad de casos $A_1, A_2, A_3, \ldots A_n, \ldots$ El caso $A_1$ es a menudo fácil; en todo caso, $A_1$ puede ser tratado con medios específicos. Una vez $A_1$ es establecida tenemos que pro ar $A_{n+1}$ asumiendo $A_n$. Una proposición $A'$ más fuerte que $A$ pued ser más fácil de probar que $A$*. En efecto, $A'$ consiste en los casos $A$ , $A'_2, \ldots, A'_n, \ldots$ Al pasar de $A$ a $A'$ hacemos más pesada la carga de la prueba: tenemos que probar la más fuerte $A'_{n+1}$ en lugar de $A_{n+1}$. Sin embargo, al mismo tiempo el apoyo de la prueba es también mayor: utilizamos la más informativa $A'_n$ en lugar de $A_n$.

La solución del ejemplo 12 nos proporciona una ilustración. Aun así, habremos hecho esta solución engorrosa por incluir los materiales tratados en el ejemplo 13, que son más convenientemente utilizados por observaciones adicionales, como un corolario.

En general, al tratar de ingeniar una prueba por inducción matemática, usted puede fracasar por dos razones opuestas. Puede fracasar por intentar probar demasiado: su $A_{n+1}$ es una carga demasiado pesada. Puede fracasar también porque intenta probar demasiado poco: su $A_n$ es un apoyo demasiado débil. En resumen, usted debe equilibrar la proposición de su teorema de modo que el apoyo sea bastante para la carga. Y así, la maquinaria de la prueba le dirige hacia un punto de vista de los hechos más equilibrado y más adaptado. Esto puede ser típico del papel de las pruebas en la edificación de la ciencia.

**15.** *Perspectiva.* Más intrincados problemas en más difíciles dominios exigen una técnica más complicada de inducción matemática y conducen a varias modificaciones de este importante método de prueba. La teoría de los grupos proporciona algunos de los ejemplos más dignos de nota. Una variante interesante es la «inducción matemática hacia atrás» o «inferencia de $n$ a $n-1$»; para un interesante ejemplo elemental cf. G. H. Hardy, J. E. Littlewood y G. Polya, *Inequalities*, págs. 17 y 20.

**16.** Siendo dado que $Q_1 = 1$ y

$$Q_{n-1}\,Q_n = \frac{1}{2^{n(n-1)}}\,\frac{n!\,(n+1)!\,(n+2)!\,\ldots\,(2n+1)!}{0!\quad 1!\quad 2!\quad \ldots\,(n-1)!}$$

para $n = 2, 3, \ldots$, encontrar y probar una expresión general para $Q_n$.

* Esta es la «paradoja del inventor»; véase *How to Solve It*, pág. 110.

**17.** *¿Son iguales cualesquiera n números?* Usted diría: No. Sin embargo, podemos intentar probar lo contrario por inducción matemática. Puede ser más atractivo, aun así, probar la aserción: «Cualesquiera $n$ muchachas tienen ojos del mismo color».

Para $n = 1$ el enunciado es obviamente (o «vacuamente») cierto. Queda pasar de $n$ a $n + 1$. Por amor a la concreción, yo pasaría de 3 a 4 y abandonaría el caso general a usted.

Déjeme presentarle cuatro muchachas cualesquiera: Ana, Berta, Carola y Dorotea, o $A$, $B$, $C$ y $D$, para abreviar. Al parecer ($n = 3$), los ojos de $A$, $B$, y $C$ son del mismo color. También los ojos de $B$, $C$ y $D$ son del mismo color, al parecer ($n = 3$). En consecuencia, los ojos de las cuatro muchachas, $A$, $B$, $C$ y $D$, deben ser del mismo color; para mayor claridad, mire el diagrama:

$$\overbrace{A,\ \underbrace{B,\ C,}\ D}$$

Esto prueba el punto para $n + 1 = 4$, y el paso de 4 a 5, por ejemplo, es, obviamente, no más difícil.

Explicar la paradoja. Usted debe intentar la aproximación experimental mirando los ojos de varias muchachas.

**18.** Si las líneas paralelas son consideradas como cortándose (en el infinito) el enunciado «Cualesquiera $n$ líneas en un plano tienen un punto común» es cierto para $n = 1$ (vacuamente) y para $n = 2$ (debido a nuestra interpretación). Construir una (paradójica) prueba por inducción matemática.

# Máximos y mínimos

*Ya que la fábrica del mundo es la más perfecta y fue construida por deseo del Creador, nada sucede en este mundo sin que alguna razón de máximos o mínimos venga a traernos luz.*—EULER.

## 1. PATRONES

Los problemas que se refieren a valores de más o de menos, esto es, los problemas de máximos y mínimos, son quizá más atractivos que otros problemas matemáticos de dificultad comparable, y esto se debe seguramente a una razón bastante primitiva. Todos nosotros tenemos problemas personales. Podemos observar que estos problemas son a menudo de una cierta clase de máximos y mínimos. Queremos obtener determinado objeto al menor precio posible, o alcanzar el mayor efecto posible con un determinado esfuerzo, o el máximo trabajo dentro de un tiempo dado, y, por supuesto, deseamos siempre correr el mínimo riesgo. Los problemas matemáticos de máximos y mínimos nos atraen porque ellos idealizan nuestros problemas cotidianos.

De la misma manera, nos inclinamos a pensar que la Naturaleza obra como a nosotros nos gustaría, obteniendo el mayor efecto con el menor esfuerzo. Los físicos logran dar forma clara y útil a ideas de esta clase; ellos describen ciertos fenómenos físicos en términos de «principios mínimos». El primer principio dinámico de esta clase (el «principio de acción mínima», que usualmente toma el nombre de Maupertuis) fue esencialmente desarrollado por Euler; sus palabras, citadas al principio de este capítulo, describen vívidamente un cierto aspecto de los problemás de máximos y mínimos que pueden haber atraído a los científicos en su siglo.

En el próximo capítulo discutiremos algunos problemas de máximos y mínimos tomándolos de la física elemental. El presente capítulo nos prepara para el próximo.

El cálculo diferencial proporciona un método general para resolver problemas de máximos y mínimos. Nosotros no utilizaremos ese método aquí. Será más instructivo en su lugar desarrollar algunos «patrones» de nuestra cosecha.

Habiendo resuelto algunos problemas con perspicacia e interés, nos encontramos con una preciosa posesión: un patrón, un modelo, que podemos imitar para resolver problemas similares. Desarrollemos este patrón si queremos ponerlo en práctica, si marca un éxito al seguirlo, si reflexionamos sobre las razones de nuestro éxito, sobre la analogía de los problemas resueltos, sobre las circunstancias relevantes que hacen a un problema accesible a esta clase de solución, etc. Desarrollando un patrón de tal tipo podemos llegar, finalmente, a un descubrimiento real. En cualquier caso, tenemos una oportunidad para obtener un conocimiento accesible y bien ordenado.

## 2. Ejemplo

*Nos son dados dos puntos y una línea recta, todos en el mismo plano y ambos puntos al mismo lado de la línea. Sobre la línea recta dada encontrar un punto desde el cual el segmento que une los dos puntos dados se ve bajo el mayor ángulo posible.*

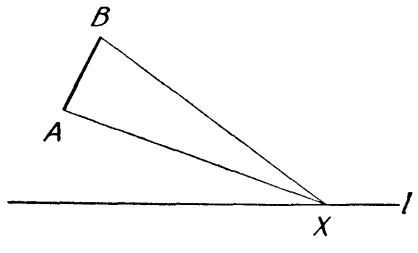



Fig. 8.1. Buscando la mejor vista.

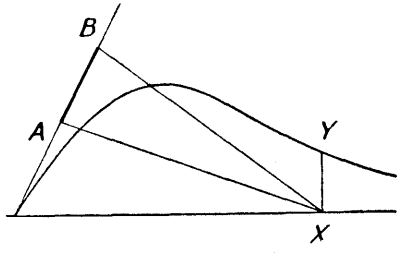



Fig. 8.2. La variación del ángulo puede parecer igual a ésta.

Este es el problema que deseamos resolver. Dibujamos una figura (figura 8.1) e introducimos una notación conveniente.

$A$ y $B$ denotan los dos puntos dados;

$l$, la línea recta dada;

$X$, un punto variable de la línea $l$.

Consideremos $< AXB$ el ángulo subtendido por el segmento dado $AB$ en el punto variable $X$. Se nos pide encontrar la posición de $X$ en la línea dada $l$, en la que dicho ángulo alcanza su máximo.

Imaginemos que $l$ es una carretera recta. Si desde algún punto de la carretera $l$ deseamos tirotear un objetivo que se extiende de $A$ a $B$ elegiremos el punto que estamos buscando; esto nos da una buena opor-

tunidad para acertar. Si tenemos la intención más pacífica de tomar, desde la carretera $l$, una instantánea de una fachada cuyas esquinas son $A$ y $B$, de nuevo tendríamos que elegir el punto que estamos buscando; esto nos da un punto de vista más amplio.

La solución de nuestro problema no es inmediata. Pero, si aún no sabemos *dónde* está el máximo, ya no nos cabe duda que *está* en alguna parte. ¿Por qué es esto tan plausible?

Podremos dar cuenta de la plausibilidad si visualizamos la variación del ángulo cuyo máximo estamos buscando. Imaginemos que paseamos por la línea $l$ y miramos al segmento $AB$. Empecemos en el punto en que la línea $l$ y la línea que pasa por $A$ y $B$ se cortan y vayamos rectamente. Al principio, el ángulo bajo el que $AB$ aparece se desvanece; después, el ángulo aumenta; aun cuando nos alejamos mucho de $AB$ disminuye otra vez, puesto que desaparece a distancia infinita *.

Entre los dos casos extremos en que el ángulo se desvanece debemos encontrar el máximo en alguna parte. Pero ¿dónde?

Esta cuestión no es fácil de contestar, aunque podríamos señalar un largo recorrido de la línea $l$ en que el máximo no se halla. Elijamos cualquier punto de la línea y llamémosle $X$. Este punto, elegido al azar, no es verosímil que sea la posición máxima que estamos buscando. ¿Cómo podríamos decidir con claridad si está en tal posición o no?

Hay una fácil y exacta observación**. Si un punto *no* está en la posición de máximo, deberá haber otro punto, al lado contrario de la posición máxima, cuyo ángulo tenga el mismo valor. ¿Hay otro punto $X'$ sobre la línea $l$ visto desde el cual el segmento $AB$ aparece bajo el mismo ángulo que desde $X$? Aquí tenemos, al fin, una cuestión que puede ser prontamente contestada: ambos $X$ y $X'$ (si hay una $X'$) deben estar en el mismo círculo que pasa por $AB$ en virtud de una familiar propiedad de los ángulos inscritos en un círculo (EUCLIDES, III, 21).

Y ahora la idea puede aparecer. Dibujemos varios círculos, que pasan por los puntos dados $A$ y $B$. Si uno de estos círculos corta la línea en dos puntos, como $X$ y $X'$ en la figura 8.3, el segmento $AB$ es visto desde ambos puntos $X$ y $X'$ bajo el mismo ángulo, pero este ángulo no es el más grande posible: un círculo que intersecta $l$ entre $X$ y $X'$ dará un ángulo aún más grande. La intersección de los círculos no puede dar lugar a engaño: el vértice del ángulo máximo es el punto en que un círculo que pasa por $A$ y $B$ *toca* la línea $l$ (el punto $M$ en la figura 8.3).

---

* Si consideramos $< AXB$ como función de la distancia medida a lo largo de la línea $l$ podemos representarlo con coordenadas rectangulares de la manera usual. La figura 8.2 da un esquema del gráfico; $< AXB$ es representado por la ordenada $XY$.

** Muy fácil si miramos la figura 8.2.

### 3. El patrón de la línea de nivel tangente

Miremos de nuevo la solución que acabamos de encontrar. ¿Qué podemos aprender de ella? ¿Qué es lo esencial? ¿Qué rasgos son capaces de una adecuada generalización?

El paso que parece más esencial después de alguna reflexión no es demasiado interesante. Este paso decisivo fue ampliar nuestro punto de vista. Salimos de la línea $l$, consideramos los valores cuantitativos que debían maximizarse (el ángulo subtendido por $AB$) en puntos del plano exteriores a $l$. Consideramos la variación de este ángulo cuando su vértice se mueve en el plano, consideramos la dependencia de este ángulo de la posición de su vértice. En resumen, concebimos este ángulo como la *función de un punto variable* (su vértice), y miramos este punto (el vértice) variando en el *plano*.

El ángulo permanece invariable cuando su vértice se mueve a lo largo de un arco del círculo que junta los puntos $A$ y $B$. Llamemos a este arco *línea de nivel*. Esta expresión subraya el punto de vista general que hemos adoptado. Las líneas a lo largo de las cuales una función de un punto variable permanece constante reciben usualmente el nombre de líneas de nivel de esa función.

Sin embargo, no olvidemos la incógnita de ese problema. Se nos pide encontrar el máximo del ángulo (de esta función de un punto variable) cuando su vértice (el punto variable) no puede moverse libremente en el plano, sino que queda restringido a una *trayectoria prescrita:* la línea $l$. ¿En qué punto de la trayectoria prescrita está el máximo buscado?

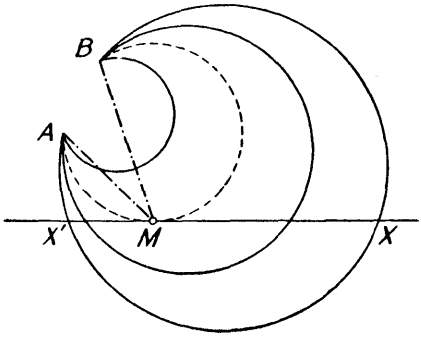



Fig. 8.3. Una línea de nivel tangente.

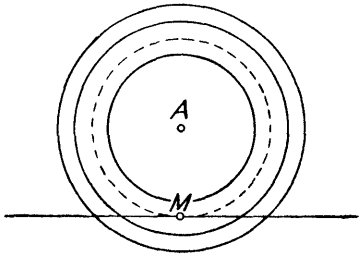



Fig. 8.4. Otra línea de nivel tangente.

Ya sabemos la contestación, pero ahora debemos comprenderla mejor y mirarla desde un punto de vista más general. Consideremos un ejemplo análogo, muy general e intuitivo.

Sabemos que las «líneas de nivel» o «curvas de nivel» están en el mapa o en el terreno (pensemos en un país montañoso) que el mapa representa. Son líneas de elevación constante; una línea de nivel conecta sobre el mapa aquellos puntos que en la superficie de la Tierra tienen la misma altura sobre el nivel del mar. Si imaginamos que el mar ha crecido 100 pies, una nueva línea de costa con bahías introduciéndose en los valles constituye un nuevo nivel del mar. Esta nueva línea costera tiene 100 pies de elevación sobre el nuevo nivel del mar. El constructor de mapas sólo dibuja unas pocas líneas de nivel, a intervalos constantes, por ejemplo, a elevaciones de 100, 200, 300, ...; aun así, podemos pensar en una línea de nivel, es decir, que *hay* una línea de nivel en cada elevación, a través de todos los puntos del terreno. La función del punto variable que es importante para el constructor de mapas, o para nosotros mismos cuando nos movemos sobre un terreno, es la elevación sobre el nivel del mar; esta elevación permanece constante a lo largo de cada línea de nivel.

Ahora tenemos aquí un problema análogo a nuestro problema discutido antes (sección 2). Paseamos a lo largo de una carretera, una trayectoria prescrita. ¿En qué punto de la carretera alcanzamos la elevación máxima?

Es muy fácil decir dónde no se alcanza. Un punto en que subimos o bajamos no es ciertamente el de máxima elevación, ni tampoco el de mínima. En tales puntos la carretera cruza de un nivel a otro: el *máximo* (o el mínimo) *no puede ser alcanzado en un punto en que la trayectoria descrita cruza una línea de nivel.*

Con esta observación esencial retornemos a nuestro ejemplo (sección 2, figuras 8.1, 8.2, 8.3). Consideremos el conjunto de la trayectoria descrita: la línea $l$ a partir de su intersección con la línea que pasa por $A$ y $B$ a distancia infinita (a la derecha). En cada uno de sus puntos esta trayectoria prescrita intersecta una línea de nivel (un arco de círculo con puntos finales en $A$ y $B$), excepto en aquel punto en que es tangente a la línea de nivel. Si el máximo está en alguna parte, debe ser en este punto: *en el punto del máximo la trayectoria pescrita es tangente a una línea de nivel.*

Esto refuerza mucho la idea general de nuestro ejemplo. Sin embargo, examinémoslo de nuevo. Apliquémoslo a un sencillo caso análogo y veamos lo que pasa. He aquí un ejemplo fácil:

*Sobre una línea recta dada encontrar el punto que está a distancia mínima de un punto dado.*

Introduzcamos una notación adecuada:

$A$ es el punto dado;

$a$ es la línea recta dada.

Se sobrentiende que el punto dado $A$ no está en la línea dada $a$. Tenemos que hallar la distancia más corta de $A$ a $a$.

Todos conocemos la solución. Imaginemos que usted está nadando en un mar en calma; en ese momento está usted en el punto $A$; la línea $a$ señala una playa uniforme. Súbitamente usted se asusta y desea alcanzar tierra firme lo antes posible. ¿Cuál es el punto más cercano a la playa? No hace falta reflexionar; hasta un perro lo sabe. Un perro o una vaca, amenazados dentro del agua, nadarían sin retraso a lo largo de las perpendiculares de $A$ a $a$.

Pero nuestro propósito no es sólo encontrar la solución, sino examinar una idea general al encontrarla. La cantidad que deseamos minimizar es la distancia a un punto variable desde otro dado, $A$. Esta distancia depende de la posición del punto variable. Las líneas de nivel de la distancia son obviamente círculos concéntricos con $A$ como centro común. La «trayectoria prescrita» es la línea recta dada $a$. El mínimo no puede ser alcanzado en un punto en que la trayectoria cruza una línea de nivel. En efecto, este mínimo es alcanzado en el (único) punto en que la trayectoria prescrita es tangente a la línea de nivel (en el punto $M$ de la figura 8.4). La distancia más corta de un punto $A$ a una línea $a$ es el radio del círculo cuyo centro $A$ es tangente a $a$ —como sabíamos desde el principio—. Aún hemos aprendido algo. La idea general aparece ahora más clara y se la puede dejar al lector para que la clarifique completamente.

Con los rasgos esenciales comunes a los problemas anteriores claramente fijos en la mente podríamos considerar problemas análogos a los que aplicar el mismo patrón de solución. En lo anterior hemos considerado un punto variable en un plano y visto el mínimo o máximo de una función de un punto tal a lo largo de una trayectoria prescrita. Pero ahora podríamos considerar un punto variable en el espacio y buscar el mínimo o máximo de una función de un punto tal a lo largo de una trayectoria prescrita, o sobre una prescrita superficie. En el plano las líneas de nivel tangente juegan un papel especial. La analogía nos impulsa a esperar que las superficies de nivel tangente juegan un papel similar en el espacio.

## 4. Ejemplos.

Discutamos dos ejemplos que pueden ser tratados por el mismo método, aunque tienen muy poco en común bajo otros aspectos.

(1) ***Encontrar la distancia mínima entre dos líneas rectas oblicuas.***

Llamemos

$a$ y $b$ las dos líneas oblicuas dadas,

$X$ un punto variable sobre $a$,

$Y$ un punto variable sobre $b$;

ver la figura 8.5. Se nos pide determinar en qué posición el segmento $XY$ será más corto.

La distancia $XY$ depende de la posición de los puntos finales, $X$ e $Y$, que son ambos variables. Hay dos puntos variables, no sólo uno, y aquí puede estar la característica dificultad del problema. Si uno de los puntos fuese dado fijo e invariable, y solamente variable el otro, el problema sería fácil. De hecho, no habría en ello nada nuevo; sería idéntico a un problema ya resuelto (sección 3).

Tomemos como *fijo por un momento* uno de los puntos originalmente variable; digamos $Y$. Entonces el segmento $XY$ está dado en el plano entre el punto fijo $Y$ y la línea dada $a$, y sólo uno de sus puntos finales, $X$, es variable, a lo largo de la línea $a$. Obviamente, $XY$ alcanza el mínimo cuando se hace perpendicular a $a$ (según la sección 3, figura 8.4).

A continuación podríamos intercambiar los papeles de los dos puntos, $X$ e $Y$, tomando $X$ como fijo e $Y$ como única variable. Obviamente, el segmento $XY$ llega a ser más corto cuando se hace perpendicular a $b$.

Sin embargo, la posición mínima de $XY$ es independiente de nuestros caprichos y de nuestra elección de papeles para $X$ e $Y$, y así somos llevados a sospechar que ambos son perpendiculares a $a$ y a $b$. Pero miremos la situación más de cerca.

De hecho, el anterior argumento nos muestra directamente dónde la posición mínima *no* puede estar (y sólo indirectamente donde estará). He aquí el resultado: una posición en que el segmento $XY$ no es perpendicular a la línea $a$ en el punto $X$ *no* es la posición mínima. En efecto, fijamos el punto $Y$ y movemos $X$ a una posición en que sea perpendicular a $a$, con lo que acortamos $XY$ (sección 3). Este razonamiento se aplica, naturalmente, tanto a $X$ como a $Y$, de donde: *la longitud del segmento $XY$ no puede ser mínima a menos que este segmento sea perpendicular tanto a* a *como a* b.

No necesitamos tomar ninguna garantía. Podemos ver de una ojeada que la perpendicular común es, en realidad, la distancia más corta. Asumamos que en la figura 8.5 el plano del dibujo es paralelo a ambas líneas dadas $a$ y $b$ ($a$ arriba, $b$ abajo). Podemos considerar cualquier punto o línea en el espacio como representado en la figura 8.5

por su proyección ortogonal. La verdadera longitud del segmento $XY$ es la hipotenusa de un triángulo rectángulo cuyo otro lado es la proyección ortogonal de $XY$ vista en la figura 8.5; el tercer lado es la distancia más corta de dos planos paralelos, uno a través de $a$, el otro a través de $b$, ambos paralelos al plano del dibujo, cuyo tercer lado es perpendicular. Por tanto, la más corta proyección de $XY$ mostrada en la figura 8.5 es $XY$ mismo. La proyección de $XY$ se reduce a un punto, su longitud a nada, y así la longitud de $XY$ es mínima si, y sólo si, $XY$ es perpendicular al plano del dibujo y también a las dos líneas $a$ y $b$.

Y así hemos verificado directamente lo que descubrimos antes por otro método.

(2) *Encontrar el máximo del área de un polígono inscrito en un círculo dado y que tiene un número dado de lados.*

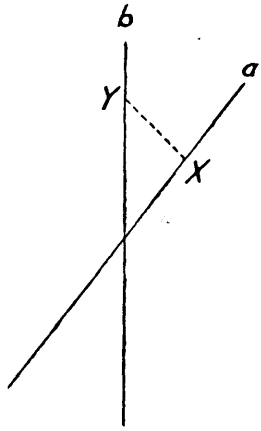



Fig. 8.5. Dos líneas oblicuas.

El círculo es dado. Sobre este círculo tenemos que elegir los $n$ vértices $U, \ldots, W, X, Y$ y $Z$ de un polígono tal que el área alcance el máximo. Como en el problema anterior (1), la principal dificultad parece ser que hay muchas variables (los vértices $U, \ldots, W, X, Y$ y $Z$). Quizá convenga utilizar el mismo método que en el anterior problema. ¿Cuál es el punto esencial de este método?

Tomemos el problema como casi resuelto. Imaginemos que hemos obtenido ya la posición deseada de todos los vértices, excepto uno, digamos $X$. Los otros $n - 1$ puntos, $U, \ldots, W, Y$ y $Z$, están ya fijos, cada uno en su posición, pero tenemos aún que elegir $X$ de modo que el área alcance un máximo. El área total consta, no obstante, de dos

partes: el polígono $U \ldots WYZ$ con $n - 1$ vértices fijos que es independiente de $X$ y el triángulo $WXY$ que depende de $X$. Traigamos la atención sobre este triángulo, que alcanzará un máximo cuando el área total alcance un máximo; ver la figura 8.6. La base $WY$ del $\Delta\ WXY$ es fija. Si el vértice $X$ se mueve a lo largo de una paralela a la base $WY$, el área permanece constante: tales paralelas a $WY$ son líneas de nivel. Tracemos la línea de nivel tangente: la tangente al círculo paralelo a $WY$. Su punto de contacto es obviamente la posición de $X$ que da el área máxima de $\Delta\ WXY$. Con $X$ en esta posición el triángulo es isósceles, $WX = XY$. Estos dos lados adyacentes deben ser iguales si el área del polígono es un máximo. Aun el mismo razonamiento se aplica a cualquier par de lados adyacentes: todos los lados deben ser iguales cuando el área máxima se ha obtenido, y así el polígono inscrito con área máxima debe ser *regular*.

## 5. El patrón de variación parcial

Comparando los dos ejemplos discutidos en la sección anterior (sección 4), reconoceremos fácilmente algunos rasgos y patrones comunes de solución. En ambos problemas buscamos el extremo (máximo o mínimo) de una cantidad, que depende de *diversos* elementos variables. En los dos casos fijamos por un momento todos los elementos originalmente variables, excepto uno, y estudiamos el efecto de la variación

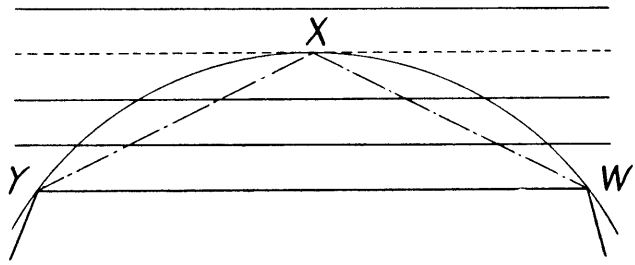



Fig. 8. 6. Triángulo de área máxima.

de este elemento solitario. La variación simultánea de todos los elementos variables, o variación *total*, no es fácil de registrar. Nosotros obtuvimos un buen resultado estudiando la variación *parcial* cuando sólo uno de los elementos varía y los otros permanecen fijos. El principio subyacente a nuestro procedimiento parecer ser: *Una función de varias*

*variables no puede alcanzar un máximo con arreglo a todas sus variables juntamente, a menos que alcance un máximo con respecto a cada variable singular.*

Este enunciado es claramente general, aunque innecesariamente restringido en un aspecto: se adhiere demasiado íntimamente a los anteriores ejemplos, en que hemos variado sólo un elemento a la vez. Sin embargo, puede ocurrir que en otros ejemplos sea ventajoso variar sólo dos elementos a la vez y fijar los otros, o quizá tres, y así sucesivamente. En tales casos, todavía podríamos hablar adecuadamente de «variación parcial». La idea general aparece ahora mucho más clara, y tras un ejemplo más el lector podrá clarificarla completamente por sí mismo.

*Una línea de longitud l es dividida en n partes. Encontrar el máximo del producto de esas n partes.*

$x_1, x_2, \ldots, x_n$ denotan las longitudes de las $n$ partes; $x_1, x_2, \ldots x_n$ son números positivos con una suma dada

$$x_1 + x_2 + \ldots + x_n = l.$$

Se pide hacer de $x_1, x_2, \ldots, x_n$ un máximo.

Examinamos primero el caso especial más sencillo: dada la suma $x_1 + x_2$ de dos cantidades positivas encontrar el máximo de su producto $x_1 x_2$. Podemos interpretar $x_1$ y $x_2$ como lados adyacentes de un rectángulo, y el problema se nos presenta en esta otra forma más atractiva: dada la longitud $L$ del perímetro de un rectángulo encontrar el máximo de su área. En efecto, la suma de los dos lados que acabamos de mencionar es la suma dada

$$x_1 + x_2 = \frac{L}{2}.$$

Hay una intuición obvia: el área alcanza un máximo cuando el rectángulo se convierte en cuadrado. Esta intuición no es difícil de verificar. Cada lado del cuadrado con perímetro $L$ es igual a

$$\frac{L}{4} = \frac{x_1 + x_2}{2}.$$

Tenemos que verificar que el área del cuadrado es mayor que el área del rectángulo, o, lo que es lo mismo, que su diferencia

$$\left(\frac{x_1 + x_2}{2}\right)^2 - x_1 x_2$$

es positiva. ¿Es esto así? Se necesita poca álgebra para ver que

$$\left(\frac{x_1 + x_2}{2}\right) - x_1 x_2 = \left(\frac{x_1 - x_2}{2}\right)^2.$$

Esta fórmula muestra toda la situación en una ojeada. El lado derecho de la ecuación es positivo, a menos que $x_1 = x_2$ y el rectángulo sea un cuadrado.

En resumen, el área de un rectángulo con un perímetro dado alcanza un máximo cuando el rectángulo es un cuadrado; el producto de dos cantidades positivas con una suma dada llega al máximo cuando estas dos cantidades son iguales.

Intentemos utilizar el caso especial que acabamos de resolver como el primer paso hacia la solución del problema general. Tomemos el problema como casi resuelto. Imaginemos que hemos obtenido ya los valores deseados de todas las partes, excepto aquellos de las dos primeras, $x_1$ y $x_2$. Luego, consideremos $x_1$ y $x_2$ como variables, y $x_3, x_4, \ldots, x_n$ como constantes. La suma de las dos partes variables es constante,

$$x_1 + x_2 = l - x_3 - x_4 - \ldots - x_n.$$

Ahora, el producto de todas las partes

$$x_1 x_2 (x_3 x_4 \ldots x_n)$$

no puede alcanzar un máximo a menos que el producto $x_1 x_2$ de las dos primeras partes alcance también un máximo. Esto requiere, no obstante, que $x_1 = x_2$. Sin embargo, no hay razón para que algún otro par de partes se conduzca de modo muy diferente. El máximo deseado del producto no puede ser alcanzado a menos que todas las cantidades con suma dada sean iguales. Citamos a Colin Maclaurin (1698-1746), a quien se debe el razonamiento anterior: «Si la línea $AB$ está dividida en cualquier número de partes $AC$, $CD$, $DE$, $EB$, el producto de todas aquellas partes multiplicadas una por otra será un *máximo* cuando las partes sean iguales entre sí».

El lector puede aprender mucho aclarando la prueba anterior. ¿Es completamente satisfactoria?

## 6. El teorema de los medios aritméticos y geométricos y sus primeras consecuencias

Volvamos a considerar el resultado de la sección anterior: Si

$$x_1 + x_2 + x_3 + \ldots + x_n = l,$$

entonces

$$x_1 x_2 x_3 \ldots x_n < \left(\frac{l}{n}\right)^n$$

a menos que $x_1 = x_2 = x_3 \ldots = x_n = l/n$. Eliminando $l$ tenemos este resultado en la forma:

$$x_1 x_2 \ldots x_n < \left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right)^n$$

o

$$\sqrt[n]{x_1 x_2 \ldots x_n} < \frac{x_1 + x_2 + \ldots + x_n}{n},$$

*a menos que todas las cantidades positivas* $x_1, x_2, \ldots, x_n$ *sean iguales;* si estas cantidades son iguales, la desigualdad se convierte en una ecuación. El lado izquierdo de la desigualdad anterior es llamado la media geométrica de $x_1, x_2, \ldots, x_n$, y el lado derecho es su media aritmética. El teorema que acabamos de establecer es llamado a veces el «teorema de las medias aritméticas y geométricas», o, más brevemente, el «teorema de las medias».

Este teorema es interesante e importante en muchos aspectos. Vale la pena mencionar que puede ser establecido de dos maneras diferentes:

*El producto de n cantidades positivas con una suma dada alcanza un máximo cuando estas cantidades son todas iguales.*

*La suma de n cantidades positivas con un producto dado alcanza un mínimo cuando estas cantidades son todas iguales.*

La primera proposición se refiere a un máximo; la segunda, al mínimo correspondiente. La derivación de la sección anterior es hecha en la primera proposición. Cambiando esta derivación sistemática podríamos llegar a la segunda proposición. Es más sencillo, sin embargo, observar que la desigualdad entre los medios da por resultado imparcial-

mente ambas proposiciones: para obtener una o la otra tenemos que considerar un lado u otro de la desigualdad. Podemos llamar a estas dos proposiciones, que son esencialmente equivalentes, *proposiciones conjugadas.*

El teorema de las medias da la solución de muchos problemas geométricos sobre máximos y mínimos. Discutimos aquí sólo un ejemplo (otros varios pueden ser encontrados al final de este capítulo).

*Dada el área de la superficie de una caja, encontrar el máximo de su volumen.*

Utilizo la palabra «caja» en vez de «paralelepípedo rectangular» porque «caja» es bastante más expresivo que el término oficial. La solución se prevé fácilmente y, una vez prevista, se reduce fácilmente al teorema de las medias como sigue. Supongamos que

$a$, $b$, $c$ denotan las longitudes de las tres aristas de la caja que salen del mismo vértice,

$S$, el área de la superficie,

$V$, el volumen.

Obviamente,

$$S = 2\,(ab + ac + bc), \qquad V = abc.$$

Observando que la suma de las tres cantidades $ab$, $ac$ y $bc$ es $S/2$ y su producto $V^2$, pensamos, naturalmente, del teorema de las medias que arrojan

$$V^2 = (abc)^2 < \left(\frac{ab + ac + bc}{3}\right)^3 = \left(\frac{S}{6}\right)^3,$$

a menos que

$$ab = ac = bc,$$

o, lo que es lo mismo,

$$a = b = c.$$

Esto es,

$$V < (S/6)^{3/2},$$

a menos que la caja sea un cubo, cuando se da la igualdad. Podemos expresar el resultado de dos formas diferentes aunque esencialmente equivalentes.

*De todas las cajas con un área de superficie dada el cubo tiene máximo volumen.*

*De todas las cajas con un volumen dado el cubo tiene la mínima área de superficie.*

Como anteriormente, podemos llamar a estas dos proposiciones proposiciones conjugadas. Como anteriormente también, una de las dos proposiciones conjugadas se refiere al máximo y la otra al mínimo.

La aplicación anterior del teorema de las medias tiene su mérito. Podemos considerarla como un patrón y recoger casos a los cuales se pueda aplicar de modo semejante el teorema de las medias.

## EJEMPLOS Y COMENTARIOS

### Primera parte

**1.** *Las distancias máximas y mínimas en geometría plana.* Encontrar la distancia mínima entre: (1) dos puntos, (2) un punto y una línea recta, (3) dos líneas rectas paralelas.

Encontrar la distancia mínima y máxima entre: (4) un punto y un círculo, (5) una línea recta y un círculo, (6) dos círculos.

La solución es obvia en todos los casos. Recordar la prueba elemental, al menos en algunos casos.

**2.** *Las distancias máximas y mínimas en geometría sólida.* Encontrar la distancia mínima entre: (1) dos puntos, (2) un punto y un plano, (3) dos planos paralelos, (4) un punto y una línea recta, (5) un plano y una línea recta paralela, (6) dos líneas rectas oblicuas.

Encontrar la distancia mínima y máxima entre: (7) un punto y una esfera, (8) un plano y una esfera, (9) una línea recta y una esfera, (10) dos esferas.

**3.** *Líneas de nivel en un plano.* Considerar la distancia de un punto variable a: (1) punto, (2) línea recta, (3) círculo, dados. ¿Qué son las líneas de nivel?

**4.** *Superficies de nivel en el espacio.* Considerar la distancia de un punto variable a: (1) punto, (2) plano, (3) línea recta, (4) esfera, dadas. ¿Qué son las superficies de nivel?

**5.** Contestar las preguntas del ejemplo 1 utilizando líneas de nivel.

**6.** Contestar las preguntas del ejemplo 2 utilizando superficies de nivel.

**7.** Dados dos lados de un triángulo encontrar el máximo del área utilizando líneas de nivel.

**8.** Dado un lado y la longitud del perímetro de un triángulo encontrar el máximo del área utilizando líneas de nivel.

**9.** Dada el área de un rectángulo encontrar el mínimo del perímetro utilizando líneas de nivel. (En un sistema rectangular de coordinadas (0, 0), ($x$, 0), (0, $y$), ($x$, $y$) son los cuatro vértices del rectángulo; utilizar la geometría analítica.)

**10.** Examinar el siguiente enunciado: «La distancia más corta desde un punto dado a una curva dada es perpendicular a la curva dada».

**11.** *El principio de la línea de paso a nivel.* Consideremos una función $f$ de un punto que varía en un plano, el máximo y el mínimo de $f$ a lo largo de una trayectoria prescrita y una línea de nivel de $f$ que separa dos regiones del plano; en una de estas regiones $f$ toma valores más altos o más bajos que sobre la línea de nivel misma.

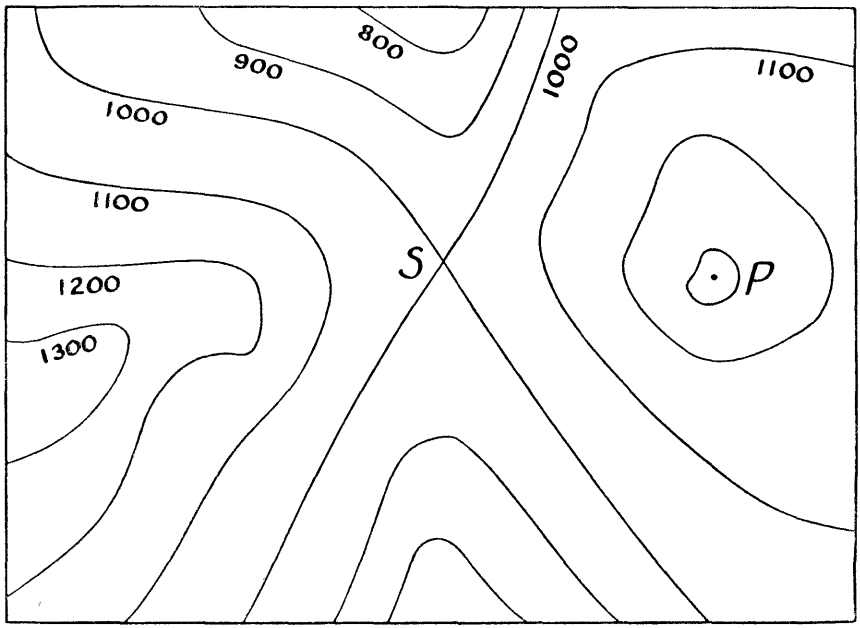



Fig. 8.7. Líneas de nivel en el mapa.

Si la trayectoria prescrita cruza la línea de nivel ni el máximo ni el mínimo de $f$ se obtiene en el punto de cruce.

**12.** El mapa de curvas de nivel de la figura 8.7 muestra un pico $P$ y un paso (o apeadero con paso horizontal tangente) $S$. Andando a lo largo de tal país ¿obtendremos necesariamente el punto más alto del camino en el punto en que éste es tangente a la línea de nivel?

**13.** $A$ y $B$ denotan dos puntos dados y $X$ un punto variable en un plano. El ángulo subtendido en $X$ por el segmento $AB$ ($< AXB$) que puede tomar cualquier valor entre 0° y 180° (incluidos los límites) es una función del punto variable $X$.

(1) Dar una descripción completa de las líneas de nivel.

(2) De dos líneas de nivel diferentes, ¿cuál corresponde al valor más alto del ángulo?

Pueden usarse las figuras 8.1 y 8.3, pero démonos cuenta que ahora es posible mirar el segmento $AB$ desde *ambos* lados.

**14.** Considerar las figuras 8.1, 8.2, 8.3, tomar $< AXB$ como en el ejemplo 13 y encontrar su mínimo a lo largo de *l*. ¿Está el resultado de acuerdo con el principio del ejemplo 11?

**15.** Dado el volumen de una caja (paralelepípedo rectangular) encontrar el mínimo del área de su superficie, utilizando variación parcial.

**16.** De todos los triángulos con un perímetro dado ¿cuál tiene el área máxima? (Ejemplo 8.)

**17.** De todos los tetraedros inscritos en una esfera dada ¿cuál tiene el volumen máximo? (¿Conoce un problema relacionado?)

**18.** Dadas $a$, $b$ y $c$, longitudes de tres aristas de un tetraedro salidas del mismo vértice, encontrar el máximo del volumen del tetraedro. (¿Conoce un problema análogo?)

**19.** Encontrar la distancia más corta entre una esfera y un cilindro. (Cilindro significa: cilindro infinito de revolución.)

**20.** Encontrar la distancia más corta entre dos cilindros con ejes oblicuos.

**21.** Examinar el siguiente enunciado: «La distancia más corta entre dos superficies dadas es perpendicular a ambas».

**22.** *El principio de variación parcial.* La función $f(X, Y, Z, \ldots)$ de varias variables $X$, $Y$, $Z$, ... obtiene su máximo para $X = A$, $Y = B$, $Z = C$, ... Luego la función $f(X, B, C, \ldots)$ de la variable única $X$ obtiene su máximo para $X = A$; y la función $f(X, Y, C, \ldots)$ de dos variables $X$ e $Y$ obtiene su máximo para $X = A$, $Y = B$, y así sucesivamente.

*Una función de varias variables no puede alcanzar un máximo con respecto a todas sus variables juntamente, a menos que alcance un máximo con arreglo a cualquier subserie de variables.*

**23.** *Existencia del extremo.* Tanto el principio de línea de nivel como el de variación parcial dan normalmente sólo «información negativa». Nos muestran directamente en qué puntos una función propuesta *f no* puede alcanzar un máximo, y de aquí podemos inferir dónde puede alcanzarlo. Que *f debe* alcanzar un máximo en algún sitio no puede derivarse de estos principios sólo. Sin embargo, la existencia del máximo puede derivarse algunas veces con una modificación del razonamiento. Por lo demás, la existencia del máximo puede a menudo derivarse de teoremas generales sobre funciones continuas de varias variables*. De cualquier modo, aunque la existencia del máximo parezca obvia desde el punto de vista intuitivo, tenemos buenas razones para esperar que algún teorema general o mecanismo especial aplicarán y probarán su existencia.

**24.** *Una modificación del patrón de variación parcial: Un proceso infinito.* Encontrar el máximo de $xyz$, dado que $x + y + z = l$.

Se entiende que $x$, $y$ y $z$ son positivos y que $l$ es dado. Este problema es un caso particular del problema de la sección 5. Siguiendo el método usado allí, hagamos fijo uno de los tres números $x$, $y$ y $z$, y cambiemos los otros dos hasta hacerlos iguales, aumentando su producto. Empezamos a partir de cualquier sistema dado $(x, y, z)$; pasamos a otro sistema $(x_1, y_1, z_1)$, realizando el cambio indicado; luego pasamos todavía a otro $(x_2, y_2, z_2)$ y de aquí a $(x_3, y_3, z_3)$, y así sucesivamente. Dejemos los tres términos sin cambiar uno tras otro: primero el término $x$, luego $y$, luego $z$, de nuevo otra vez $x$, luego $y$, luego $z$, una vez más $x$, y así sucesivamente. Tendremos así el conjunto

$$x_1 = x, y_1 = z_1 = \frac{y + z}{2},$$

$$y_2 = y_1, z_2 = x_2 = \frac{z_1 + x_1}{2},$$

$$z_3 = z_2, x_3 = y_3 = \frac{x_2 + y_2}{2},$$

$$x_4 = x_3, y_4 = z_4 = \frac{y_3 + z_3}{2},$$

. . . . . .

* Una función de varias variables, continua en un conjunto cerrado, alcanza allí sus límites más altos y bajos. Véase G. H. HARDY, *Pure Mathematics*, pág. 194, teorema 2.

Cada paso deja incambiada la suma, pero incrementa el producto:

$$x + y + z = x_1 + y_1 + z_1 = x_2 + y_2 + z_2 = \ldots$$

$$xyz < x_1 y_1 z_1 < x_2 y_2 z_2 < \ldots$$

Hemos asumido que $y \neq z$ y que $x_1 \neq z_1$. (Este es un caso no-excepcional; en el caso excepcional alcanzamos nuestro objetivo mucho más fácilmente.) Esperamos, naturalmente, que los tres números $x_n$, $y_n$ y $z_n$ se hacen menos diferentes unos de otros a medida que $n$ aumenta. *Si* podemos probar, finalmente, que

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n$$

podemos inferir de inmediato que

$$xyz < \lim_{n\to\infty} x_n y_n z_n = (l/3)^3.$$

Obtenemos así este resultado a considerable gasto, aunque *sin asumir la existencia del máximo* desde el comienzo.

Probar que $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n$.

**25.** *Otra modificación del patrón de modificación parcial: Un proceso finito.* Nos referimos todavía al problema del ejemplo 24, pero utilizamos ahora una modificación más complicada del método de la sección 5.

Tenemos $l = 3A$; así, $A$ es la media aritmética de $x$, $y$ y $z$, y tenemos

$$(x - A) + (y - A) + (z - A) = 0.$$

Puede suceder que $x = y = z$. Si no se da ese caso, una de las diferencias de la izquierda de la ecuación debe ser negativa y la obra positiva. Elijamos la notación de modo que

$$y < A < z.$$

Pasamos ahora del sistema $(x, y, z)$ al $(x', y', z')$, poniendo

$$x' = x, \quad y' = A, \quad z' = y + (z - A);$$

donde dejamos la primera cantidad sin cambiar. Luego,

$$x + y + z = x' + y' + z'$$

e

$$y'z' - yz = A(y + z - A) - yz$$
$$= (A - y)(z - A) > 0,$$

de modo que

$$xyz < x'y'z'.$$

Puede suceder que $x' = y' = z'$. Si no se da este caso pasamos de $(x', y', z')$ a $(x'', y'', z'')$, poniendo

$$y'' = y', \qquad z'' = x'' = \frac{z' + x'}{2}$$

que da

$$x'' = y'' = z'' = A$$

y de nuevo aumenta el producto (como sabemos desde la sección 5) de manera que

$$xyz < x'y'z' < x''y''z'' = A^3.$$
$$= \left(\frac{x + y + z}{3}\right)^3.$$

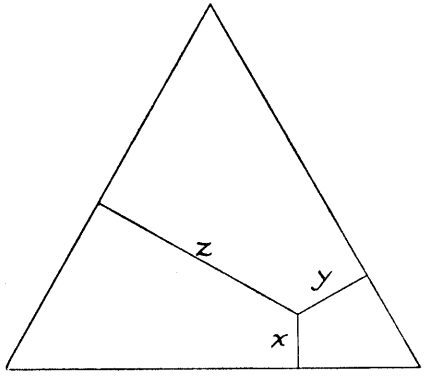



Fig. 8.8. Coordenadas triangulares.

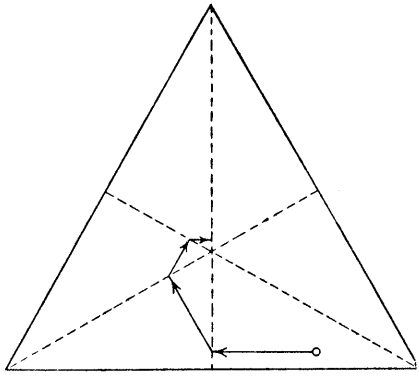


Fig. 8.9. Pasos sucesivos de aproximación al centro.

Hemos probado el resultado deseado, *sin asumir la existencia del máximo* y *sin considerar límites*.

Mediante una conveniente extensión de este procedimiento probar el teorema de las medias (sección 6) en general para $n$ cantidades.

**26.** *Confrontación gráfica.* Sea $P$ un punto en el interior de un triángulo equilátero de altura $l$ y sean $x$, $y$ y $z$ las distancias a $P$ desde los tres lados del triángulo; ver la figura 8.8. Luego,

$$x + y + z = l.$$

(¿Por qué?) Los números $x$, $y$ y $z$ son las *coordenadas triangulares* del punto $P$. Cualquier sistema de tres números positivos $x$, $y$ y $z$ con suma $l$ pueden ser interpretadas como las coordenadas triangulares de un único punto determinado dentro del triángulo.

La secuencia

$$(x, y, z), \quad (x_1, y_1, z_1) \quad (x_2, y_2, z_2)$$

considerada en el ejemplo 24 se representa con una secuencia de puntos en la figura 8.9. Los segmentos que unen los puntos consecutivos son paralelos a los varios lados del triángulo sucesivamente, al primero, al segundo y al tercero, respectivamente, y luego otra vez al primero, y así sucesivamente. (¿Por qué?) El procedimiento del ejemplo 25 es representado por tres puntos y dos segmentos. (¿Cómo?)

**27.** Considerar el argumento de la sección 4 (2) y modificarlo tomando primero el ejemplo 24 como modelo, y después el ejemplo 25.

**28.** Una condición necesaria para un valor máximo o mínimo de una función $f(x, y, z)$ en el punto $(a, b, c)$ es la de las derivadas parciales

$$\frac{\delta f}{\delta x}, \frac{\delta f}{\delta y}, \frac{\delta f}{\delta z}$$

desaparecen para $x = a, y = b, z = c$.

La prueba usual de este teorema ejemplifica uno de nuestros modelos. ¿Cuál?

**29.** Establecer la condición necesaria y conocida (en términos de derivadas parciales) para un valor máximo o mínimo de la función $f(x, y)$ bajo la condición lateral (o subsidiaria) de que $x$ e $y$ estén ligados por la ecuación $g(x, y) = 0$. Explicar la conexión con el modelo de la línea de nivel tangente.

**30.** Reexaminar los casos mencionados en la solución del ejemplo 12 a la luz de la condición mencionada en el ejemplo 29. ¿Hay alguna contradicción?

**31.** Establecer la condición necesaria y conocida para un valor máximo o mínimo de una función $f(x, y, z)$ bajo la condición lateral de que $g(x, y, z) = 0$. Explicar la conexión con el modelo de la superficie de nivel tangente.

**32.** Establecer la condición necesaria y conocida para el valor máximo o mínimo de una función $f(x, y, z)$ bajo las dos simultáneas y laterales condiciones de que $g(x, y, z) = 0$ y $h(x, y, z) = 0$. Explicar la conexión con el modelo de la superficie de nivel tangente.

## Segunda parte

La terminología y notación utilizadas en lo que sigue son explicadas en el ejemplo 33, que debe leerse en primer lugar.

**33.** *Polígonos y poliedros. Area y perímetro. Volumen y superficie.* En nuestro trato con los polígonos usaremos la siguiente notación la mayor parte del tiempo:

$A$ para el área, y

$L$ para la longitud del perímetro.

En lo referente a poliedros,

$V$ denota el volumen, y

$S$ el área de la superficie.

Discutiremos problemas de máximos y mínimos referentes a $A$ y $L$, o $V$ y $S$. Tales problemas eran conocidos de los antiguos griegos *. Nosotros discutiremos principalmente problemas tratados por Simón Lhuilier y Jacob Steiner **. Las desigualdades algebraicas elementales, especialmente el teorema de las medias (sección 6), serán muy útiles para resolver la mayoría de los problemas siguientes.

La mayoría de las veces estos problemas tratan sólo con los polígonos más sencillos (triángulos y cuadriláteros) y los más sencillos poliedros (prismas y pirámides). Tendremos que aprender algunos términos poco corrientes.

Dos pirámides que se levantan a lados opuestos de su base común forman en conjunto una *pirámide doble*. Si la base tiene $n$ lados, la pirámide doble tiene $2n$ caras, $n + 2$ vértices y $3n$ aristas. La base *no* es una cara de la pirámide doble.

---

* Pappus, *Collectiones*, libro V.

** Simón Lhuilier, *Polygonométrie et Abrégé d'Isopérimétrie élémentaire*, Génova, 1789; J. Steiner, *Gesammelte Werke*, vol. 2, págs. 177-308.

Si todas las caras son perpendiculares a la base, llamaremos al prisma, *prisma recto.*

Si la base de una pirámide está circunscrita en un círculo y la altura coincide en la base con el centro del círculo, la llamaremos *pirámide recta.*

Si las dos pirámides formando una doble pirámide son rectas y simétricas una a otra con respecto a su base común, la llamaremos *pirámide doble recta.*

Si un prisma, pirámide o doble pirámide no es «recto», le llamaremos «oblicuo». Entre los cinco sólidos regulares hay un solo prisma, una sola pirámide y una sola doble pirámide: el cubo, el tetraedro y el octaedro, respectivamente. Cada uno de éstos es un sólido «recto» de su clase.

Consideraremos también cilindros, conos y dobles conos; si no hay indicación en contrario, las bases se supone que son círculos.

**34.** *Prisma recto con base cuadrada.* De todos los prismas rectos con base cuadrada que tienen un volumen dado el cubo es el de superficie mínima.

Probar este caso particular de un teorema ya probado (sección 6, ejemplo 15) directamente, utilizando el teorema de las medias.

Podemos ser tentados a proceder como sigue: $V$, $S$, $x$ e $y$ denotan, respectivamente, el volumen, el área de la superficie, el lado de la base y la altura del prisma. Luego,

$$V = x^2y, \qquad S = 2x^2 + 4xy.$$

Aplicando el teorema de las medias, obtenemos

$$(S/2)^2 = [(2x^2 + 4xy)/2]^2 \geqq 2x^2 \cdot 4xy = 8x^3y.$$

Pero esto no parece tener relación útil con $V = x^2y$; el teorema de las medias no parece ser aplicable.

No obstante, esto parece haber sido una irreflexiva, atropellada y poco profesional aplicación del teorema. Probemos otra vez. [¿Cuál es la conclusión deseada?]

**35.** *Cilindro recto.* Observar que, de todos los prismas considerados en el ejemplo 34, sólo el cubo está circunscrito por una esfera, y probar: De todos los cilindros rectos que tienen un volumen dado el cilindro circunscrito por una esfera tiene el mínimo de superficie. [¿Cuál es la conclusión deseada?]

**36.** *Prisma recto general.* Tenemos un prisma recto con el volumen y la forma de la base (aunque no el tamaño) dados. Cuando el

área de la superficie es un mínimo, ¿qué fracción de ésta es el área de la base? [¿Conoce un problema relacionado?]

**37.** *Pirámide doble recta con base cuadrada.* Probar: De todas las pirámides dobles rectas con base cuadrada que tienen un volumen dado el octaedro regular tiene el mínimo de superficie.

**38.** *Cono recto doble.* Observar que la esfera inscrita toca cada cara del octaedro regular en su centro, el cual divide la altura de la cara del radio 1 : 2, y probar. De todos los conos dobles rectos que tienen un volumen dado el mínimo de la superficie es alcanzado por el cono doble, cuyas generatrices están divididas en el radio 1 : 2 por los puntos de contacto con la esfera inscrita.

**39.** *Pirámide doble recta general.* De una pirámide doble recta se nos da el volumen y la forma (aunque no el tamaño) de la base. Cuando el área de la superficie es un mínimo, ¿qué fracción de ésta es el área de la base?

**40.** Dado el área de un triángulo encontrar el mínimo de su perímetro. [¿Podría usted predecir el resultado? Si desea probar el teorema de las medias puede necesitar la expresión del área en términos de los lados.]

**41.** Dado el área de un cuadrilátero encontrar el mínimo de su perímetro. [¿Podría predecir el resultado? Llamemos $a$, $b$, $c$ y $d$ los lados del cuadrilátero, $\varepsilon$ la suma de dos ángulos opuestos y expresar el área $A$ en términos de $a$, $b$, $c$, $d$, y $\varepsilon$. Esto es una generalización del problema resuelto por la fórmula de Herón.]

**42.** Un prisma recto y otro oblicuo tienen el mismo volumen y la misma base. Luego el prisma recto tiene la superficie más pequeña.

Una pirámide recta y una pirámide oblicua tienen el mismo volumen y la misma base. Luego la pirámide recta tiene la superficie más pequeña.

Una pirámide doble recta y una pirámide doble oblicua tienen el mismo volumen y la misma base. Luego, la pirámide doble recta tiene la superficie más pequeña.

En estos tres enunciados las bases de los dos sólidos comparados concuerdan en forma y tamaño. (Los volúmenes, por supuesto, concuerdan sólo en tamaño.)

Elegir el enunciado que parezca más asequible de los tres y probarlo.

**43.** *Aplicando la geometría al álgebra.* Probar: Si $u_1$, $u_2$, ... $u_n$, $v_1$, $v_2$, ... $v_n$ son números reales,

$$\sqrt{u_1^2 + v_1^2} + \sqrt{u_2^2 + v_2^2} + \ldots + \sqrt{u_n^2 + v_n^2}$$
$$\geqq \sqrt{(u_1 + u_2 + \ldots + u_n)^2 + (v_1 + v_2 + \ldots + v_n)^2}$$

y la desigualdad se alcanza si, y sólo si,

$$u_1 : v_1 = u_2 : v_2 = \ldots = u_n : v_n.$$

[Considerar $n + 1$ puntos $P_0$, $P_1$, $P_2$, ... $P_n$ en un sistema de coordenadas rectangulares y la longitud de la línea $P_0P_1P_2 \ldots P_n$.]

**44.** Probar la desigualdad del ejemplo 43 independientemente de consideraciones geométricas. (En la prueba geométrica de la desigualdad el caso especial guía es $n = 2$.)

**45.** *Aplicando el álgebra a la geometría.* Probar: De todos los triángulos con base y área dada, el isósceles tiene el perímetro mínimo. [Ejemplo 43.]

**46.** $V$, $S$, $A$ y $L$ denotan el volumen, el área de la superficie total, el área de la base y la longitud del perímetro de la base de una pirámide $P$, respectivamente. $V_0$, $S_0$, $A_0$ y $L_0$ denominan las cantidades correspondientes a otra pirámide $P_0$. Suponiendo que

$$V = V_0, \quad A = A_0, \quad L \geqq L_0$$

y que $P_0$ es una pirámide recta, probar que

$$S \geqq S_0.$$

La igualdad es obtenida si, y sólo si, $L = L_0$ y $P$ es también una pirámide recta. [Ejemplo 43.]

**47.** $V$, $S$, $A$ y $L$ denotan el volumen, el área de la superficie, el área de la base y el perímetro de la base de una pirámide doble $D$, respectivamente. $V_0$, $S_0$, $A_0$ y $L_0$ denominan las cantidades correspondientes a otra pirámide doble $D_0$. Suponiendo que

$$V = V_0, \quad A = A_0, \quad L \geqq L_0$$

y que $D_0$ es una pirámide doble recta, probar que

$$S \geqq S_0.$$

La igualdad es obtenida si, y sólo si, $L = L_0$ y $D$ es también una pirámide doble recta. (Ejemplos 45, 46.]

**48.** Probar: De todos los prismas cuadriléteros con un volumen dado el cubo tiene la superficie mínima. [Comparar con el ejemplo 34; ¿qué enunciado es más fuerte?]

**49.** Probar: De todas las pirámides dobles cuadrangulares con un volumen dado el octadedro regular tiene la superficie mínima. [Comparar con el ejemplo 37; ¿qué enunciado es más fuerte.]

**50.** Probar: De todas las pirámides triangulares con un volumen dado el tetraedro regular tiene la superficie mínima.

**51.** *Pirámide recta con base cuadrada.* Probar: De todas las pirámides rectas con base cuadrada que tienen un volumen dado aquella cuya base es 1/4 de la superficie total tiene la superficie mínima.

**52.** *Cono recto.* De todos los conos rectos que tienen un volumen dado el cono cuya base es 1/4 de la superficie total tiene el mínimo de superficie.

**53.** *Pirámide recta general.* Tenemos una pirámide recta con el volumen y la forma (aunque no el tamaño) de la base dados. Cuando el área de la superficie es un mínimo ¿qué fracción de ésta es el área de la base? [¿Puede señalar un caso particular?]

**54** Volviendo la mirada a nuestros distintos ejemplos de prismas, pirámides y dobles pirámides, observar sus relaciones mutuas y ordenarlas en una tabla de modo que la analogía de los *resultados* se haga visible. Apuntar los vacíos que espera llenar con nuevos resultados.

**55.** *La caja sin tapa.* Dado $S_5$, la suma de las áreas de las cinco caras de una caja, encontrar el máximo del volumen $V$. [¿Conoce algún problema relacionado? ¿Podría utilizarse aquí su resultado o método?]

**56.** *La artesa.* Dado $S_4$, la suma de las áreas de cuatro caras de un prisma triangular recto (la cara perdida es una *lateral*), encontrar el máximo del volumen $V$.

**57.** *Un fragmento.* En un prisma recto con base triangular se nos da $S_3$, la suma de las áreas de tres caras mutuamente adyacentes (es decir, de dos caras laterales y una base). Mostrar que estas tres caras tienen el mismo área y son perpendiculares unas a otras cuando el volumen $V$ alcanza su máximo. (Un fragmento —¿de qué?—.)

**58.** Dada el área de un sector de un círculo encontrar el valor del ángulo en el centro cuando el perímetro es un mínimo.

**59.** En un triángulo se nos da un área y un ángulo. Encontrar el mínimo: (1) de la suma de los dos lados incluyendo el ángulo dado, (2) del lado opuesto al ángulo dado, (3) de todo el perímetro.

**60.** Se nos da en posición un ángulo y un punto en el plano del ángulo, dentro del ángulo. Una línear ecta variable, que pasa a través del punto dado, recorta un triángulo desde el ángulo. Encontrar el mínimo del área de este triángulo.

**61.** Dada $E$, la suma de las longitudes de las doce aristas de una caja, encontrar el máximo: (1) de su volumen $V$, (2) de su superficie $S$.

**62.** *Un problema de correos.* Encontrar el máximo del volumen de una caja, dado que la longitud y el contorno combinados no exceden de $l$ pulgadas.

**63.** *Un problema de Kepler.* Dado $d$, la distancia desde el centro de un generador en forma de cilindro recto hasta el punto más lejano de ese cilindro; encontrar el máximo del volumen del cilindro.

# Matemática física

*La ciencia física no sólo nos da [a los matemáticos] una oportunidad para resolver problemas, sino que nos ayuda también a encontrar los medios de resolverlos, y esto de dos maneras: anticipando la solución y sugiriendo apropiadas líneas de argumentación.*—HENRI POINCARÉ *.

## 1. INTERPRETACIÓN ÓPTICA

Los problemas matemáticos son a menudo inspirados por la naturaleza o, más bien, por nuestra interpretación de la naturaleza, la ciencia física. También la solución de un problema matemático puede ser inspirado por la naturaleza: la física nos proporciona claves que por nuestra propia cuenta hubiéramos tenido muy poca oportunidad de apropiarnos. Nuestra perspectiva sería demasiado estrecha sin discutir problemas matemáticos sugeridos por la investigación física y resueltos con ayuda de la interpretación física. He aquí un primer problema muy sencillo de esta clase.

(1) *La naturaleza sugiere un problema.* La línea recta es la distancia más corta entre dos puntos dados. La luz, viajando a través del aire de un punto a otro, elige dicha distancia; al menos así nos lo muestra la experiencia diaria. Pero ¿qué sucede cuando la luz viaja indirectamente de un punto a otro, sufriendo una reflexión o un espejo interpuesto? ¿Elige entonces la luz también la distancia más corta? ¿Cuál es la distancia más corta en estas circunstancias? Consideraciones sobre la propagación de la luz nos conducen al siguiente problema puramente geométrico:

*Nos son dados dos puntos y una línea recta, todo en el mismo plano, y los dos puntos al mismo lado de la línea. Sobre una línea recta dada encontrar un punto tal que la suma de sus distancias desde los dos puntos dados sea un mínimo.* (Véase figura 9.1.)

$A$ y $B$ denotan los dos puntos dados;

$l$, la línea recta dada;

$X$, un punto variable de la línea $l$.

* *La valeur de la science*, pág. 152.

Consideremos $AX + XB$, la suma de dos distancias o, lo que es lo mismo, la longitud de la distancia de $A$ a $X$ y de aquí a $B$. Se nos pide encontrar la posición de $X$ sobre la línea $l$ para que la longitud de esta trayectoria sea un mínimo.

Hemos visto ya un problema semejante anteriormente (sección 8.2, figuras 8.1, 8.2, 8.3). En efecto, ambos problemas tienen exactamente los mismos datos e incluso la incógnita es de la misma naturaleza: Aquí, como allí, buscamos la posición de un punto sobre una línea dada para la cual se obtiene un cierto extremo. Los dos problemas difieren sólo en la naturaleza de este extremo: Aquí buscamos minimizar la suma de las dos líneas; allí tratábamos de maximizar el ángulo formado por estas dos líneas.

Aun así, los dos problemas están tan íntimamente relacionados que es natural intentar el mismo método. Para resolver el problema de la sección 8.2 utilizamos líneas de nivel; volvamos a utilizarlas aquí.

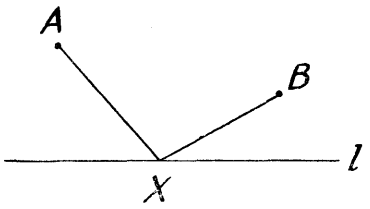



Fig. 9.1. ¿Qué trayectoria es la más corta?

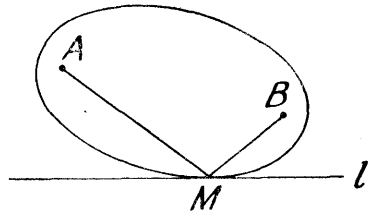



Fig. 9.2. Una línea de nivel tangente.

Consideremos que el punto $X$ no es el límite de la trayectoria prescrita, sino que es libre de moverse dentro del plano. Si la cantidad $AX + XB$ (que deseamos minimizar) tiene un valor constante, ¿cómo puede $X$ moverse? A lo largo de una elipse con focos $A$ y $B$. Por tanto, las líneas de nivel son elipses «confocales», es decir, elipses con el mismo foco (los puntos dados $A$ y $B$). *El mínimo deseado es alcanzado en el punto de contacto de la trayectoria prescrita $l$ con una elipse cuyos focos son los puntos dados $A$ y $B$* (ver figura 9.2).

(2) *La naturaleza sugiere una solución.* Hemos encontrado la verdadera solución. Sin embargo, a menos que conozcamos algunas propiedades geométricas de la elipse, la solución no nos será muy útil. Hagamos un nuevo comienzo y busquemos una solución más informativa.

Visualicemos la situación física que sugiere nuestro problema. El punto $A$ es una fuente de luz; el punto $B$, el ojo de un observador, y $l$ señala la posición de una superficie plana que refleja la luz; podemos pensar en la superficie horizontal de una piscina tranquila (perpendicu-

lar al plano de la figura 9.1, que la intersecta en la línea $l$). La línea quebrada $AXB$ representa, si el punto $X$ es correctamente elegido, la trayectoria de la luz; conocemos esta trayectoria muy bien, por experiencia. Sospechamos que la longitud de la línea quebrada $AXV$ es un mínimo cuando representa la trayectoria real de la luz reflejada.

El ojo que está en la posición $B$ verá reflejada en la piscina la imagen de $A$. El rayo de luz que percibe no viene directamente de $A$, sino que parece provenir de algún punto bajo la superficie de la piscina. ¿De qué punto? Del punto $A^*$, la imagen del objeto $A$ reflejada en el espejo, simétrica a $A$ con respecto a la línea $l$.

¡Introduzcamos el punto $A^*$, sugerido por la experiencia física, dentro de la figura! Este punto $A^*$ cambia la faz del problema. Vemos una multitud de nuevas relaciones (figura 9.3), que procedemos a ordenar y explotar rápidamente. Obviamente,

$$AX = A^*X.$$

($A^*X$ es la imagen en el espejo de $AX$. También se puede argumentar desde la correspondencia de los triángulos $\Delta ACX$, $\triangle A^*CX$; la línea $l$ es el bisector perpendicular del segmento $AA^*$.) Por tanto,

$$AX + XB = A^*X + XB.$$

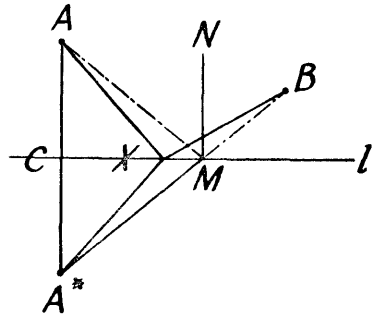



Fig. 9.3. Una solución más informativa.

Ambos lados de esta ecuación están minimizados por la misma posición de $X$. Sin embargo, el lado derecho es obviamente un *mínimo cuando $A^*$, $X$ y $B$ están sobre la misma línea recta.* La líneá recta es la más corta.

Esta es la solución (ver figura 9.3). El punto $M$, posición mínima de $X$, se obtiene como intersección de la línea $l$ y de la línea que une $A^*$

y $B$. Evidentemente, $AM$ y $MB$ forman el mismo ángulo con $l$. Introduciendo la línea $MN$, normal a $l$ (paralela a $A^*A$), vemos que

$$< AMN = < BMN.$$

La igualdad de estos dos ángulos caracteriza la trayectoria más corta. Sin embargo, la misma igualdad:

$$\text{ángulo de incidencia} = \text{ángulo de reflexión},$$

caracteriza la trayectoria real de la luz, como sabemos por experiencia. Además, en efecto, el rayo de luz reflejado sigue el camino más corto posible entre el objeto y el ojo. Este descubrimiento se debe a Herón de Alejandría.

(3) *Comparando dos soluciones.* A menudo es útil volver a mirar la solución completa. En el caso presente será doblemente útil, pues tenemos dos soluciones que podemos comparar una con otra (bajo (1) y (2)). Ambos métodos de resolver el problema (figuras 9.2 y 9.3) deben arrojar el mismo resultado (imaginar las dos figuras superpuestas). Podemos obtener el punto $M$, la solución de nuestro problema mínimo, por medio de una elipse tangente a $l$, o por medio de dos rayos igualmente inclinados a $l$. Sin embargo, estas dos construcciones deben concordar cualquiera que pueda ser la posición relativa de los datos (los puntos $A$ y $B$ y la línea $l$). La concordancia de las dos soluciones implica una propiedad geométrica de la elipse: *Las dos líneas rectas que unen los dos focos de una elipse con un punto de la periferia de la elipse están igualmente inclinadas a la tangente de la elipse en el punto en que se encuentran.*

Si concebimos la elipse como un espejo y tenemos en cuenta la ley de reflexión que acabamos de discutir, podemos volver a formular la propiedad geométrica desde una interpretación óptica intuitiva: *un rayo de luz que viene de uno de los focos de un espejo elíptico se refleja en el otro foco.*

(4) *Una aplicación.* Tan sencillo como es, el descubrimiento de Herón merece un puesto en la historia de la ciencia. Es el primer ejemplo del uso de un principio mínimo en la descripción de fenómenos físicos. Un sugestivo ejemplo de interrelaciones entre teorías físicas y matemáticas. Principios mínimos mucho más generales se han descubierto, y teorías físicas y matemáticas han sido interrelacionadas en mayor escala después de Herón, pero los primeros y más sencillos ejemplos son los más impresionantes en varios aspectos.

Volviendo la vista al éxito logrado en (2), preguntamos: ¿Se puede utilizar? ¿Se puede utilizar el resultado? ¿Se puede utilizar el método? De hecho, hay varias posibilidades. Podríamos examinar la reflexión de la luz en un espejo curvado, o reflexiones sucesivas en una serie de espejos planos, o combinar el resultado con métodos que aprendimos antes, y así sucesivamente.

Vamos a examinar aquí sólo un ejemplo: el problema del «centro de tráfico». Tres ciudades tratan de construir tres carreteras con un centro de tráfico común, el cual será elegido de forma que el costo total de construcción de la carretera sea el mínimo. Si tomamos todo esto lo más sencillamente posible tenemos el siguiente problema puramente geométrico: *Dados tres puntos encontrar un cuarto de forma que la suma de sus distancias desde los tres puntos dados sea la mínima.*

Supongamos que $A$, $B$ y $C$ denotan los tres puntos dados (ciudades) y $X$ un punto variable del plano determinado por $A$, $B$ y $C$. Buscamos el mínimo de $AX + BX + CX$.

Este problema parece relacionarse con el problema de Herón. Consideraremos ambos problemas unidos, estableciendo la más estrecha relación posible entre ellos. Si por un momento tomamos la distancia $CX$ como fija (pongamos $= r$), la relación parece muy estrecha: aquí como allí tenemos que encontrar el mínimo de $AX + BX$, la suma de las distancias de un punto variable desde dos puntos fijos. La diferencia es que $X$ tiene que moverse a lo largo de un círculo en este caso (con radio $r$ y centro $C$), y a lo largo de una línea recta en aquel otro. El primer problema trataba de la reflexión en un espejo plano, el presente trata sobre la reflexión en un espejo circular. Confiemos en la luz: ésta es lo bastante lista como para encontrar el camino más corto desde $A$ al espejo circular y desde aquí a $B$. No obstante, la luz se mueve de manera que el ángulo de incidencia es igual al ángulo de reflexión. Por tanto, en la posición mínima deseada $< AXB$ debe tener como bisectriz la línea recta que pasa por $C$ y $X$ (ver figura 9.4). Por el principio de variación parcial y la simetría de la situación, $< AXC$ y $< BXC$ deben tener bisectrices de modo semejante. Las tres líneas rectas, que unen $X$ a $A$, $B$ y $C$, dividen el plano en seis ángulos cuyo vértice común es $X$. Centrando nuestra atención en los pares de ángulos verticales de la figura 9.5 vemos fácilmente que los seis ángulos son iguales, y, por tanto, cada uno de ellos es igual a 60°. *Las tres carreteras que salen del centro del tráfico están igualmente inclinadas cada una a la otra;* el ángulo entre dos cualesquiera de ellas es de 120°.

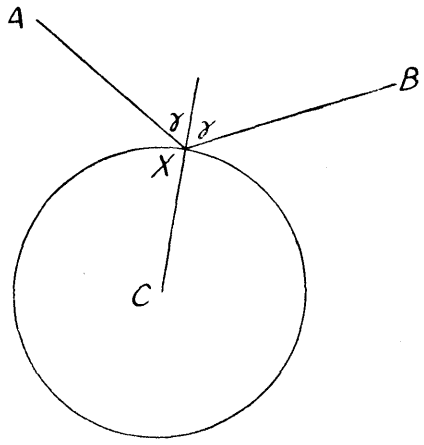



Fig. 9.4. Centro de tráfico y espejo circular.

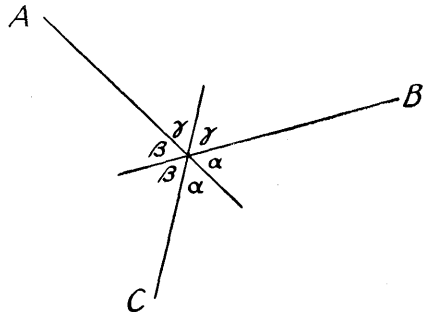



Fig. 9.5. El centro de tráfico.

(Si recordamos que el método de variación parcial que hemos usado está sujeto a ciertas limitaciones podemos volver a un examen crítico de nuestra solución.)

## 2. Interpretación mecánica.

Los problemas matemáticos y sus soluciones pueden ser sugeridos por cualquier sección de nuestra experiencia, bien sea óptica, mecánica u otra clase de fenómenos. Ahora vamos a discutir cómo sencillos principios mecánicos pueden ayudarnos a descubrir la solución.

(1) *Una cuerda cuyos extremos están fijos pasa a través de un pesado anillo; encontrar la posición de equilibrio.*

Se entiende que la cuerda es perfectamente flexible e inextensible. Su peso es insignificante, el anillo corre a lo largo de la cuerda sin ninguna fricción, y las dimensiones de este último son tan pequeñas que pueden ser consideradas como un punto matemático.

$A$ y $B$ denotan los puntos finales y fijos de la cuerda, y $X$ cualquier posición del anillo. La cuerda forma la línea quebrada $AXB$ de la figura 9.6.

El problema propuesto se puede resolver por dos métodos diferentes.

Primero, el anillo cuelga tan bajo como sea posible. (En efecto, el anillo es pesado; «quiere» acercarse, tanto como sea posible, al suelo o al centro de la Tierra.) Ambas partes de la cuerda inextensible, $AX$ y $BX$, están tirantes, y de esta forma el anillo, deslizándose a lo largo

de la cuerda, describe una elipse con focos $A$ y $B$. Obviamente, la posición de equilibrio está en el punto más bajo $M$ de la elipse donde la tangente es *horizontal.*

Segundo, las fuerzas que actúan sobre el punto $M$ de la cuerda deben estar en equilibrio. El peso del anillo y las tensiones de la cuerda actúan sobre el punto $M$. Las tensiones de ambas partes de la cuerda, $MA$ y $MB$, son iguales y dirigidas a lo largo de la cuerda hacia $A$ y $B$, respectivamente.

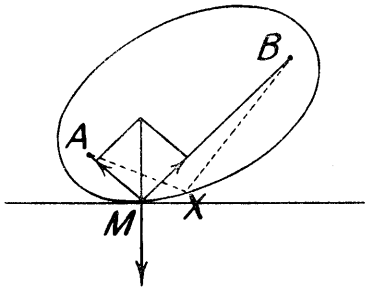



Fig. 9.6. Dos condiciones de equilibrio.

Su resultante biseciona el $< AMB$ y, siendo opuesta al peso del anillo, es *vertical.*

No obstante, las dos soluciones están de acuerdo. Por tanto, las líneas $MA$ y $MB$, igualmente inclinadas a la normal vertical de la elipse, están también igualmente inclinadas a su tangente horizontal: *Las dos líneas rectas que unen los focos de una elipse a cualquier punto de M de la periferia están igualmente inclinadas a la tangente en el punto M.* (Para guardar la longitud $AB$, aunque cambiando su ángulo de inclinación a la horizontal, traemos $M$ a una posición deseada sobre la mitad de la elipse.)

Derivamos un primer resultado (sección (3)) por un nuevo método que puede ser capaz de posteriores aplicaciones.

(2) Parece que tenemos un exceso de conocimientos. Sin necesidad de haber aprendido mucha mecánica sabemos bastante no sólo para encontrar una solución al problema mecánico propuesto, sino incluso dos soluciones, basadas en dos principios diferentes. La comparación de estas dos soluciones nos conducen a un interesante hecho geométrico. ¿Podríamos verter algo más de este exceso de conocimientos mecánicos en otros canales?

Con un poco de suerte podemos imaginar un mecanismo para resolver el problema del centro de tráfico considerado arriba (sección 1 (4)): Tres poleas giran alrededor de clavos fijos en una pared vertical en los

puntos $A$, $B$ y $C$; ver la figura 9.7. Tres cuerdas, $XAP$, $XBQ$ y $XCR$, en la figura 9.7 pasan sobre las poleas en $A$, $B$ y $C$, respectivamente. En el punto común $X$ las tres cuerdas atadas en uno de sus extremos y en el otro soportan pesos, $P$, $Q$ y $R$, respectivamente. Estos pesos son iguales. Nuestro problema es encontrar la posición de equilibrio.

Por supuesto, este problema puede comprenderse con las usuales simplificaciones: Las cuerdas son flexibles e inextensibles; la fricción, el peso de las cuerdas y las dimensiones de las poleas son insignificantes (las poleas son consideradas como puntos). Como en (1), podemos resolver el problema por métodos diferentes.

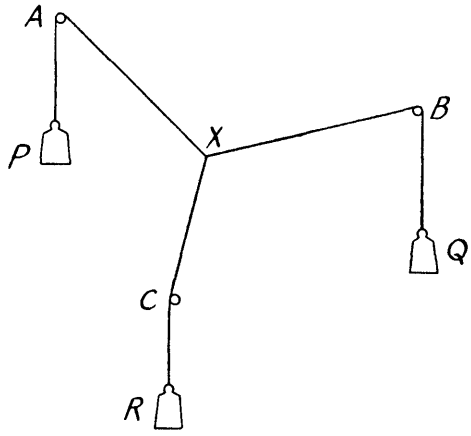



Fig. 9.7. El centro de tráfico por mecanismo mecánico.

Primero, los tres pesos pueden colgar tan bajo como sea posible. Es decir, la suma de sus distancias desde un nivel horizontal dado (el suelo) debe ser un mínimo. (Esto es, la energía potencial del sistema debe ser un mínimo; recordemos que los tres pesos son iguales.) Por tanto, $AP + BQ + CR$ deben ser un máximo. Y, puesto que la longitud de cada cuerda es invariable, $AX + BX + CX$ debe ser un mínimo, y así nuestro problema se hace idéntico con el problema del centro de tráfico de la sección 1 (4), figuras 9.4, 9.5.

Segundo, las fuerzas que actúan en el punto $X$ deben estar en equilibrio. Los tres pesos iguales, cada uno en su cuerda, tiran con igual fuerza y estas fuerzas son transmitidas sin disminución por las poleas sin fricción. Las tres fuerzas iguales que actúan en $X$ a lo largo de las líneas $XA$, $XB$ y $XC$, respectivamente, deben estar en equilibrio. Obviamente, por simetría, deben estar igualmente inclinadas unas a otras; el ángulo entre dos cualesquiera de las tres cuerdas que se reúnen

en el punto $X$ es de 120°. (El triángulo formado por las tres fuerzas es equilátero, sus ángulos exteriores son de 120°.)

Esto confirma la solución de la sección 1 (4). (Por otro lado, la interpretación mecánica puede enfatizar la necesidad de alguna restricción referente a la configuración de los tres puntos $A$, $B$ y $C$.)

## 3. Reinterpretación

Un palo, sumergido hasta la mitad en el agua, parece que está torcido. De aquí concluimos que la luz que sigue un curso recto tanto en el agua como en el aire sufre un brusco cambio de dirección al pasar del agua al aire. Este es el fenómeno de la refracción, fenómeno aparentemente más complicado y difícil de comprender que la reflexión. La ley de refracción fue descubierta por Snellius (hacia 1621) y publicada por Descartes, después de los infructuosos esfuerzos de Kepler y otros. Aun después, Fermat (1601-1655) recogió el hilo de las ideas iniciado por Herón.

La luz, que va desde un objeto $A$ situado debajo del agua a un ojo $B$, que está sobre el agua, describe una línea quebrada con un punto angular sobre la superficie que separa el aire del agua; ver figura 9.8. La línea recta es, sin embargo, el camino más corto entre $A$ y $B$, y así, la luz, en su transición de un medio a otro, deja de obedecer al principio de Herón. Esto es desalentador; no nos gusta admitir que una regla sencilla que sirvió en dos casos (propagación directa y reflexión) fracase en un tercero (refracción). Fermat encontró un recurso. El estaba familiarizado con la idea de que la luz tarda un tiempo para viajar de un punto a otro, de que viaja con una cierta velocidad (finita); de hecho, Galileo propuso un método para medir la velocidad de la luz. Quizá la luz que se propaga con cierta velocidad a través del aire se propaga con otra velocidad a través del agua; tal diferencia de velocidades podría explicar quizá el fenómeno de la refracción. La luz, en tanto se propaga a una velocidad constante, eligiendo el camino más corto, elige también el *más rápido*. Si la velocidad depende del medio atravesado el camino más corto no es necesariamente el más rápido. Quizá la luz elige *siempre* el curso más rápido, incluso cuando va del agua al aire.

Este curso de ideas conduce a un claro problema de mínimos (ver figura 9.8): *Dados dos puntos, $A$ y $B$, una línea recta $l$, que separa $A$ de $B$, y dos velocidades, $u$ y $v$, encontrar el tiempo mínimo necesitado para viajar de $A$ a $B$; se supone que se viaja de $A$ a $l$ a la velocidad $u$, y de $l$ a $B$ a la velocidad $v$.* Como es obvio, lo más rápido es seguir una línea

recta desde $A$ hasta un cierto punto $X$ situado sobre $l$, y otra línea recta desde $X$ hasta $B$. El problema consiste, esencialmente, en encontrar el punto $X$. Ahora bien, en el movimiento uniforme el tiempo es igual a la distancia partida por la velocidad. Por tanto, el tiempo que se tarda desde $A$ a $X$ y desde aquí a $B$ es

$$\frac{AX}{u} + \frac{XB}{v}.$$

Esta cantidad alcanzará un mínimo eligiendo convenientemente el punto $X$ situado en la línea $l$. Tenemos que encontrar $X$ teniendo los datos de $A$, $B$, $u$, $v$ y $l$. Resolver este problema sin hacer uso del cálculo diferencial no es demasiado fácil. Fermat lo resolvió inventando un método que al final le condujo a dicho cálculo. Nosotros seguiremos la dirección dada por los ejemplos de la sección anterior. Con un poco de suerte conseguiremos imaginar un mecanismo que resuelva el problema de mínimos propuesto; véase figura 9.9.

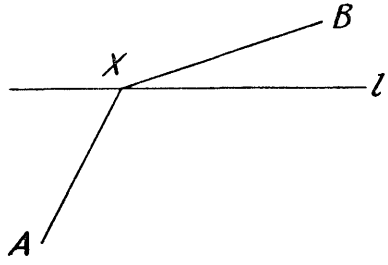



Fig. 9.8. Refracción.

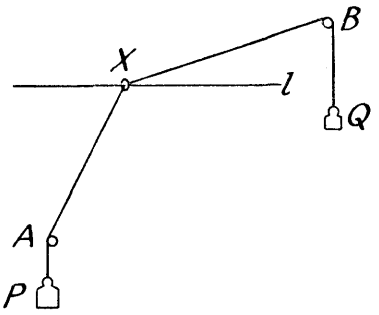



Fig. 9.9. Refracción por ingenio mecánico.

Un anillo $X$ se desliza a lo largo de una barra horizontal fija. Dos cuerdas $XAP$ y $XBQ$ están atadas al anillo $X$. Cada una de estas cuerdas pasa por una polea (en $A$ y $B$, respectivamente) y llevan un peso en su otro extremo (en $P$ y $Q$, especialmente). La cuestión principal es elegir los pesos, que no pueden ser iguales; si lo fueran, la línea $AXB$ sería recta en la posición de equilibrio (al menos esto parece lo plausible) y así $AXB$ no representaría adecuadamente la trayectoria de la luz refractada. Vamos a dejar la elección de los pesos, pero introduzcamos una conveniente notación. Llamemos $p$ al peso situado en el extremo $P$ de la primera cuerda y $q$ al peso situado en el extremo $Q$ de la segunda cuerda. Y ahora tenemos que encontrar la posición de equilibrio (se dan por supuestas las simplificaciones usuales: la barra es completamente

inflexible, las cuerdas perfectamente flexibles e inextensibles; no tenemos en cuenta la fricción, el peso y rigidez de las fuerzas, las dimensiones de las poleas y del anillo). Como en la sección 2, resolvemos el problema por dos métodos diferentes.

Primero, los dos pesos deben colgar juntos tan bajo como sea posible. (Esto es, la energía potencial del sistema ha de ser mínima.) Esto implica que

$$AP \cdot p + BQ \cdot q$$

debe ser un máximo. Por tanto, puesto que la longitud de cada cuerda es invariable,

$$AX \cdot p + XB \cdot q$$

debe ser un mínimo.

Este problema está muy próximo al de Fermat, pero no es exactamente el mismo. Sin embargo, los dos problemas, el óptico y el mecánico, coinciden matemáticamente si elegimos

$$p = 1/u, \qquad q = 1/v.$$

Entonces el problema de equilibrio (figura 9.9) requiere, lo mismo que el problema del viaje más rápido de Fermat, que

$$\frac{AX}{u} + \frac{XB}{v}$$

sea un mínimo. Esto hemos encontrado observando el equilibrio del sistema mecánico de la figura 9.9 desde un primer punto de vista.

Segundo, las fuerzas que actúan en el punto $X$ deben estar en equilibrio. La tracción de los pesos es transmitida sin disminuir por las poleas sin fricción: dos fuerzas de magnitud $1/u$ y $1/v$, respectivamente,

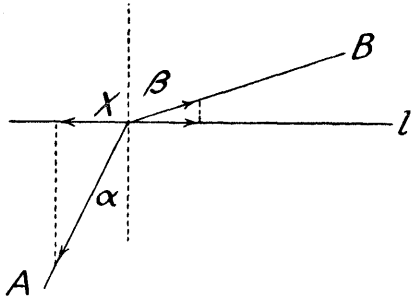



Fig. 9.10. La ley de refracción.

actúan sobre el anillo tirando de él en la dirección de la cuerda respectiva. No pueden moverlo en la dirección vertical porque la barra $l$ que pasa a través del anillo es completamente rígida (hay una reacción vertical de cantidad ilimitada debida a la barra). Sin embargo, los *componentes horizontales* de las dos tracciones que están en dirección opuesta deben anularse, deben ser *iguales en magnitud.* Para expresar esta relación introducimos los ángulos $\alpha$ y $\beta$ entre la vertical que pasa por el punto $X$ y las dos cuerdas; véase figura 9.10. La igualdad de los componentes horizontales se expresa por

$$\frac{1}{u} \operatorname{sen} \alpha = \frac{1}{v} \operatorname{sen} \beta$$

o

$$\frac{\operatorname{sen} \alpha}{\operatorname{sen} \beta} = \frac{u}{v}.$$

Esta es la *condición de mínimo.*

Ahora volvamos a la interpretación óptica. El ángulo $\alpha$, entre el rayo entrante y la normal a la superficie refractante, se llama ángulo de incidencia, y $\beta$, entre el rayo que sale y la normal, ángulo de refracción. La razón $u/v$ de las velocidades depende de los dos medios, agua y aire, pero no de circunstancias geométricas como la situación de los puntos $A$ y $B$. Por tanto, la condición de mínimo requiere que los *senos de los ángulos de incidencia y refracción estén en una relación constante que depende sólo de los dos medios* (esa relación o razón es llamada hoy índice de refracción). El «principio del menor tiempo» de Fermat conduce a la ley de refracción de Snellius, confirmada por numerosísimas observaciones.

Hemos reconstruido lo mejor que hemos podido el nacimiento de un importante descubrimiento. El procedimiento de solución (que hemos usado en lugar del de Fermat) es también digno de nota. Nuestro problema tuvo desde el comienzo una interpretación física (óptica). Sin embargo, al objeto de resolverlo, inventamos *otra* interpretación física (mecánica). Nuestra solución lo fue *por reinterpretación.* Tales soluciones pueden revelar nuevas analogías entre diferentes fenómenos físicos y tienen una peculiar cualidad artística.

## 4. El descubrimiento por Jean Bernouille de la braquistocrona

Un punto material se pone en marcha desde un punto en descanso $A$ y se desliza sin fricción a lo largo de un plano inclinado hasta el punto más bajo $B$. El punto material en descanso podría también oscilar de $A$ a $B$ a lo largo de un arco circular, como la bola de un péndulo. ¿Qué movimiento tarda menos, el que va a través de la línea recta o el que va a través del arco circular? Galileo pensó que el último era más rápido. Jean Bernouille imaginó una curva arbitraria en el plano vertical, que pasa por $A$ y $B$ conectando estos dos puntos. Hay infinitas de tales curvas, pero él quería encontrar la curva que hace del tiempo de descenso un mínimo; esta curva recibió el nombre de «curva del más rápido descenso» o «braquistocrona». Por nuestra parte, deseamos comprender la solución maravillosamente imaginativa que Jean Bernouille dio a este problema.

Colocamos una curva arbitraria que desciende de $A$ a $B$ en un sistema de coordenadas; véase la figura 9.11. Elegimos $A$ como origen,

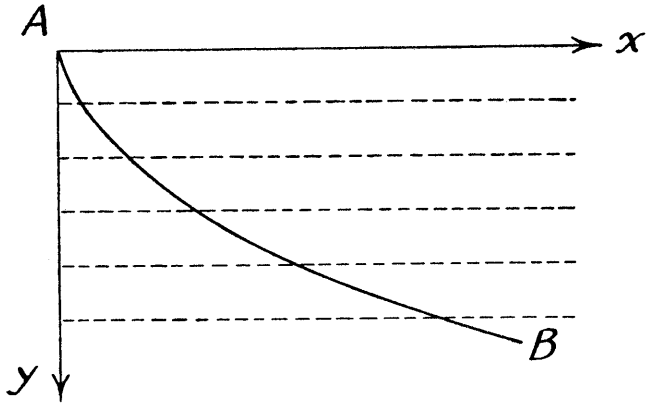



Fig. 9.11. Trayectoria de un punto material.

el eje de las $x$ es horizontal, mientras el de las $y$ se dirige verticalmente hacia abajo. Por un momento fijémonos en el punto material que al deslizarse por la curva pasa por un cierto punto $(x, y)$ con cierta velocidad $v$. Entonces tenemos la relación

$$v^2/2 = gy,$$

que era perfectamente familiar a Bernouille; hoy la derivamos de la conservación de la energía. Es decir, cualquiera que sea la trayectoria

del descenso, $v$, la velocidad obtenida depende sólo de $y$, la profundidad del descenso:

$$v = (2gy)^{1}/_{2}. \tag{1}$$

¿Qué significa esto? Intentemos ver intuitivamente la significación de este hecho básico.

Dibujamos líneas horizontales (ver figura 9.11) que dividen el plano en capas por las que desciende el punto material. Este punto descendente cruza estas capas una tras otra. Su velocidad no depende de la trayectoria que sigue, sino de las capas que cruza; su velocidad varía de capa a capa. ¿Dónde hemos visto una situación semejante? Cuando la luz del Sol viene a nosotros tiene que cruzar varias capas de aire con densidad diferente; por tanto, la velocidad de la luz varía de capa a capa. El problema mecánico propuesto admite una *interpretación óptica.*

Ahora veamos la figura 9.11 en un nuevo contexto. Miremos esta figura como si representase un medio ópticamente heterogéneo. Este medio está estratificado con estratos de cualidad diferente; la velocidad de la luz en la capa horizontal a profundidad $y$ es $(2gy)^{1}/_{2}$. La luz que cruza este medio de $A$ a $B$ (desde uno de los puntos dados a otro) *podría* viajar a través de varias curvas. Sin embargo, la luz elige el camino más rápido; ella viaja, en realidad, a lo largo de la curva que hace del tiempo de viaje un mínimo. Por tanto, *¡la trayectoria real de la luz, que atraviesa un medio estratificado y heterogéneo de A a B, es la braquistocrona!* Sin embargo, la trayectoria real de la luz está gobernada por la ley de refracción de Snellius: la solución aparece súbitamente al alcance. La reinterpretación imaginativa de Jean Bernouille nos hace accesible un problema que parecía enteramente nuevo e inaccesible.

Aún nos queda, a pesar de todo, algo que hacer, aunque requiere mucha menos originalidad. Al objeto de hacer la ley de Snellius aplicable en su forma familiar (que hemos discutido en la sección 3 precedente) hemos vuelto a cambiar nuestra interpretación de la figura 9.11, ligeramente: la velocidad $v$ no varía continuamente con $y$ en grados infinitesimales, sino discontinuamente en grados pequeños. Imaginemos varias capas horizontales de materia transparente (algunos platos de cristal), cada uno de ellos algo diferente ópticamente de sus vecinos. Sean $v, v', v'', v''', \ldots$, la velocidad de la luz en capas sucesivas, y la luz que las cruza sucesivamente incluye los ángulos $\alpha, \alpha', \alpha'', \alpha''', \ldots$ con la vertical, respectivamente; véase la figura 9.12. Por la ley de Snellius (véase sección 3),

$$\frac{\operatorname{sen}\,\alpha}{v} = \frac{\operatorname{sen}\,\alpha'}{v'} = \frac{\operatorname{sen}\,\alpha''}{v''} = \frac{\operatorname{sen}\,\alpha'''}{v'''} = \ldots$$

Ahora podemos volver desde el medio consistente en delgados platos al medio estratificado en que $v$ varía continuamente con la profundidad. (Hagamos que los platos se hagan infinitamente delgados.) Veamos entonces que

$$\frac{\operatorname{sen}\,\alpha}{v} = \text{const} \tag{2}$$

a lo largo de la trayectoria de la luz.

Sea $\beta$ el ángulo formado por la tangente a la curva con la horizontal. Luego,

$$\alpha + \beta = 90°; \quad \tan\beta = dy/dx = y'$$

y así,

$$\operatorname{sen}\,\alpha = \cos\beta = (1 + y^{12})^{-1/2} \tag{3}$$

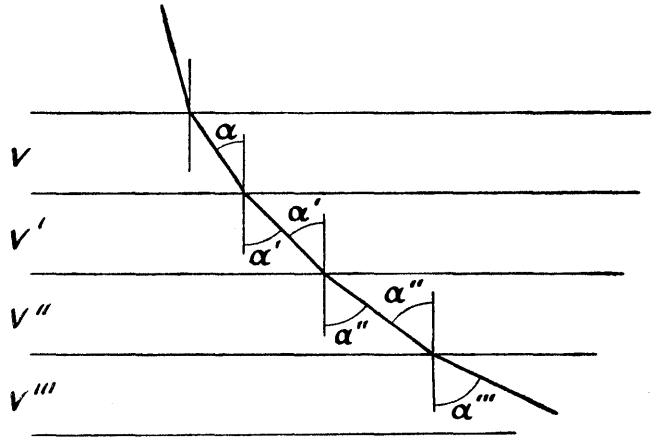



Fig. 9.12. Trayectoria de la luz.

Combinamos las ecuaciones (1), (2) y (3) (derivados de la mecánica, de la óptica y del cálculo, respectivamente), introducimos una notación conveniente para la constante de (2) y obtenemos

$$y\,(1 + y'^2) = c;$$

donde $c$ es una constante positiva. Hemos obtenido una ecuación diferencial de primer orden para la braquistocrona. Encontrar las curvas que

satisfacen tal ecuación era un problema familiar a Bernouille. No necesitamos entrar aquí en detalles (ver, no obstante, el ejemplo 31): la braquistocrona determinada por la ecuación diferencial se convierte en una *cicloide*. (La cicloide es descrita por un punto de la circunferencia de un círculo que gira una línea recta; en nuestro caso, la línea recta es el eje de las $x$ y el giro procede de arriba abajo: el círculo gira bajo el eje de las $x$.)

Observemos, no obstante, que podemos ver intuitivamente, sin recurrir a las fórmulas, que la ley de Snellius implica una ecuación diferencial. En efecto, esta ley determina las direcciones de los sucesivos elementos de la trayectoria representada en la figura 9.12, y esto es, precisamente, lo que hace una ecuación diferencial.

La solución de Jean Bernouille al problema de la braquistocrona que hemos discutido aquí tiene una peculiar cualidad artística. Mirando la figura 9.11 ó la 9.12 podemos ver intuitivamente la idea clave de la solución. Si podemos ver esta idea claramente, sin esfuerzo, anticipando lo que implica, nos daremos cuenta que hay un verdadero trabajo de arte ante nosotros.

La idea clave de la solución de Jean Bernouille es, por supuesto, reinterpretación. La figura geométrica (figura 9.11 ó 9.12) es concebida sucesivamente en dos interpretaciones diferentes, vista en dos diferentes «contextos»: primero en un contexto matemático, luego en uno óptico. ¿Consiste todo descubrimiento en un contacto inesperado, y la subsecuente interpretación, de dos contextos diferentes?

## 5. El descubrimiento por Arquímedes del cálculo integral

Uno de los más grandes descubrimientos matemáticos de todos los tiempos fue orientado por la intuición física. Me refiero al descubrimiento de Arquímedes referente a esa rama de la ciencia que hoy llamamos cálculo integral. Arquímedes encontró el área del segmento parabólico, el volumen de la esfera y una docena de resultados semejantes mediante un método uniforme en que la idea de equilibrio juega un importante papel. Como él mismo dice de sí mismo, «investigó algunos problemas matemáticos por medios mecánicos» *.

Si deseamos comprender algo de la obra de Arquímedes debemos

* El *Método*, de Arquímedes, editado por Thomas L. Heath, Cambridge, 1912. Cf. pág. 13. Este librito será citado por *Método* en las siguientes notas a pie de página. Véase también *Oeuvres complètes d'Archimède*, traducidas por P. Ver Eecke, páginas 474-519.

conocer algo sobre el nivel de conocimiento desde el que empezó.

La geometría de los griegos alcanzó su cumbre en tiempo de Arquímedes; Eudoxio y Euclides fueron sus predecesores, Apolonio su contemporáneo. Hemos mencionado unos pocos puntos específicos que pueden haber influido en el descubrimiento de Arquímedes.

Como el mismo Arquímedes refiere, Demócrito encontró el volumen del cono; estableció que éste es un tercio del volumen de un cilindro con la misma base y la misma altura. No sabemos nada sobre el método de Demócrito, aunque hay buenas razones para sospechar que consideró lo que hoy llamaríamos una sección transversal del cono paralela a su base *.

Eudoxio fue el primero en probar el enunciado de Demócrito. Al probar este y resultados semejantes Eudoxio inventó su «método de agotamiento» y estableció un nivel de rigor para la matemática griega.

Nos hemos dado cuenta que los griegos conocían, en cierto sentido, «la geometría de coordenadas». Estas fueron usadas para manejar lugares de un plano considerando las distancias de un punto movible desde dos ejes fijos de referencia. Si la suma de los cuadrados de estas distancias es constante y los ejes de referencia son perpendiculares uno a otro, el lugar es un círculo —esta proposición pertenece a la geometría de coordenadas, pero todavía no a la geometría analítica—. La geometría analítica empieza en el momento en que expresamos la relación mencionada en símbolos algebraicos como

$$x^2 + y^2 = a^2.$$

La mecánica de los griegos nunca llegó a la altura de su geometría y se inició mucho más tarde. Si discutimos con Aristóteles y otros acerca de lo que tiene de valor, podemos afirmar que la mecánica como ciencia empieza con Arquímedes. El descubrió, como todo el mundo sabe, la ley de los cuerpos flotantes; igualmente descubrió el principio de la palanca y las principales propiedades del centro de gravedad, que pronto utilizaremos.

Ahora estamos preparados para discutir el ejemplo más espectacular de toda la obra de Arquímedes; deseamos encontrar con este método el *volumen de la esfera*. Arquímedes considera la esfera como engendrada por un círculo que gira y considera el círculo como un lugar caracterizado por una relación entre las distancias de un punto variable desde dos fijos ejes rectangulares de referencia. Escrita en notación moderna, esta relación es

$$x^2 + y^2 = 2ax,$$

* Cf. *Método*, págs. 10-11.

la ecuación de un círculo de radio $a$ que toca al eje de las $y$ en el origen. Ver la figura 9.13, que sólo difiere ligeramente de la figura original de Arquímedes; el círculo girando sobre el eje de las $x$ engendra una esfera. Creo que el uso de notación moderna no falsea la idea de Arquímedes. Por el contrario, me parece que esta notación es muy sugestiva. Sugiere motivos que pueden conducirnos a la idea de Arquímedes hoy y que no son quizá demasiado diferentes de los motivos que condujeron al mismo Arquímedes a su descubrimiento.

En la ecuación del círculo tenemos el término $y^2$. Observar que $\pi y^2$ es el área de una *sección transversal variable de la esfera.* Sin embargo, Demócrito encontró el volumen del cono examinando la variación de su sección transversal. Esto nos conduce a volver a escribir la ecuación del círculo en la forma

$$\pi x^2 + \pi y^2 = \pi 2\, ax.$$

Ahora podemos interpretar $\pi x^2$ como la variable sección de un cono, engendrado por rotación de la línea $y = x$ sobre el eje de las $x$; ver figura 9.13.

Esto sugiere buscar una interpretación análoga del término $\pi 2ax$. Si no vemos tal interpretación, aún podemos volver a escribir la ecuación en otras formas, y así tenemos una oportunidad de acertar con la forma

$$2a\,(\pi\, y^2 + \pi x^2) = x\pi\,(2a)^2. \qquad \text{(A)}$$

Hay mucho concentrado en esta ecuación (A). Mirándola observamos varias longitudes y áreas que surgen de ella, y *disponiéndolas convenientemente en la figura* podemos ser testigos del nacimiento de una gran idea, que nacerá de la íntima unión de la fórmula (A) con la figura 9.13.

Observemos las áreas de los tres discos circulares, $\pi y^2$, $\pi x^2$ y $\pi(2a)^2$. Los tres círculos son las intersecciones del mismo plano con tres sólidos de rotación. El plano es perpendicular al eje de las $x$ y está a la distancia $x$ del origen $O$. Los tres sólidos de rotación son una esfera, un cono y un cilindro. Están descritos por las tres líneas de las ecuaciones que son (A), $y = x$ e $y = 2a$, respectivamente, cuando la parte del lado derecho de la figura 9.13 gira sobre el eje $x$. El cono y el cilindro tienen la misma base y la misma altura. El radio de la base y la altura común tiene la misma longitud $2a$. El cono del vértice está en el origen $O$.

Arquímedes trata de modo distinto los discos cuyas áreas aparecen a los distintos lados de la ecuación (A). El deja el disco de radio $2a$, sección del cilindro, en su posición original, a la distancia $x$ del origen. Quita, por el contrario, los discos de radios $x$ e $y$, secciones del cono y la

esfera, respectivamente, de su posición original y los traslada al punto $H$ del eje de las $x$ con abscisa $-2a$. Estos discos de radios $y$ y $x$ cuelgan con su centro verticalmente bajo $H$, suspendidos por una cuerda de insignificante peso; ver figura 9.13. (La cuerda es un añadido de muy poca monta a la figura original de Arquímedes.)

Consideremos el eje de las $x$ como una *palanca*, una barra rígida de peso desdeñable, y el origen $O$ como una *cuña* (fulcro) o punto de apoyo. La ecuación (A) se refiere a *momentos*. (Un momento es el producto del peso y el brazo de la palanca.) La ecuación (A) expresa que el momento de los dos discos del lado izquierdo es igual al momento del disco del lado derecho, y así, por ley mecánica descubierta por Arquímedes, la palanca está en equilibrio.

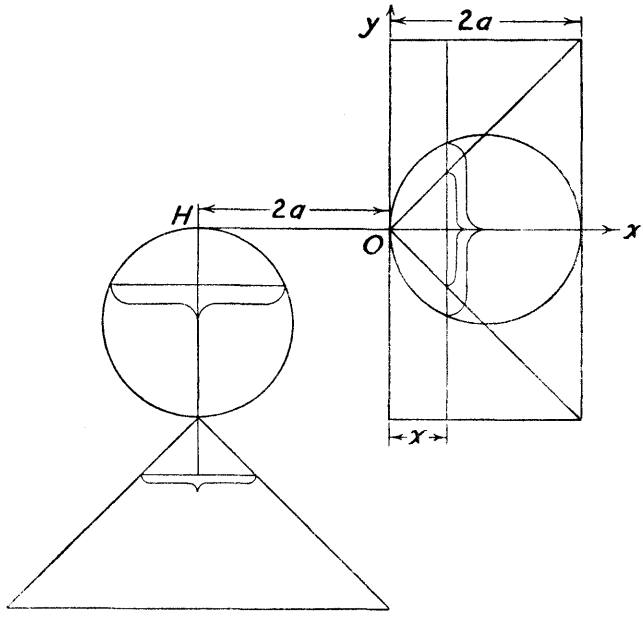



Fig. 9.13. El nacimiento del cálculo integral.

Como $x$ varía de 0 a $2a$, obtenemos todas las secciones transversales del cilindro; estas secciones *constituyen* el cilindro. A cada sección transversal del cilindro corresponden dos suspendidas del punto $H$ y estas secciones *constituyen* la esfera y el cono, respectivamente. *Como sus correspondientes secciones tranversales, la esfera y el cono, que cuelgan de H, están en equilibrio con el cilindro.* Por tanto, según la ley mecánica de Arquímedes, los momentos deben ser iguales. Llamemos $V$ el volumen de la esfera, recordemos la expresión para el volumen del cono

(debida a Demócrito) y también el volumen del cilindro y la obvia localización de su centro de gravedad. Pasando de los momentos de las secciones transversales a los momentos de los sólidos correspondientes nos vemos conducidos de la ecuación (A) a

$$2a\left(V + \frac{\pi(2a)^2 2a}{3}\right) = \alpha\pi(2a)^2 2a, \qquad \text{(B)}$$

que no da, realmente *,

$$V = \frac{4\pi a^3}{3}.$$

Volviendo la vista a lo anterior, vemos que el paso decisivo es el de (A) a (B), de las secciones transversales constituyentes a los cuerpos sólidos completos. Sin embargo, este paso es asumido heurísticamente, aunque no justificado lógicamente. Es plausible, incluso muy plausible, pero no demostrativo. Es una intuición, no una prueba. Y Arquímedes, representante de la gran tradición del rigor matemático griego, lo sabía muy bien: «El hecho a que hemos llegado no es, en realidad, demostrado por el argumento utilizado, pero este argumento es una especie de indicación de que la conclusión es cierta» **. Esta intuición, no obstante, es de gran alcance. La idea va mucho más lejos de lo que requiere el problema entre manos, y tiene un horizonte inmensamente más grande. El paso de (A) a (B), de las secciones transversales al sólido total, es, en lenguaje más moderno, la transición de la parte infinitesimal a la cantidad total, de lo diferencial a lo integral. Esta transición es un gran comienzo, y Arquímedes, que fue un hombre bastante grande para verse a sí mismo con perspectiva histórica, lo sabía muy bien: «Estoy persuadido de que este método será de no poca utilidad a los matemáticos. Preveo que este método, una vez comprendido, será usado para descubrir otros teoremas que no se me han ocurrido a mí, por otros matemáticos, vivientes o aún no nacidos» ***.

---

* He presentado esta derivación varias veces en mis clases y una de ellas recibí un cumplido, del que estoy orgulloso. Después de mi usual: "¿Hay alguna pregunta?" al fin de la derivación, un muchacho me preguntó: "¿Quién pagó a Arquímedes esta investigación?". Confieso que no fui muy rápido en contestar: «En aquellos tiempos tales investigaciones estaban sólo costeadas por Urania, la musa de la Ciencia».

** *Método*, pág. 17.

*** *Método*, pág. 14.

## EJEMPLOS Y COMENTARIOS

**1.** Se nos da en un plano un punto $P$ y dos líneas $l$ y $m$ que se intersectan, aunque ninguna de ellas pasa por $P$. Sea $Y$ un punto variable sobre $l$ y $Z$ un punto variable sobre $m$. Determinar $Y$ y $Z$ de modo que el perímetro de $\Delta PYZ$ sea un mínimo.

Dar dos soluciones, una por consideraciones físicas y otra geométricas.

**2.** Tres círculos en un plano, exteriores uno a otro, se nos dan en posición. Encontrar el triángulo con perímetro que tiene un vértice sobre cada círculo.

Dar dos interpretaciones físicas distintas.

**3.** *Triángulo con perímetro mínimo inscrito en un triángulo dado.* Dado $\Delta ABC$. Encontrar tres puntos $X$, $Y$ y $Z$ sobre los lados $BC$, $CA$ y $AB$ del triángulo, respectivamente, tal que el perímetro de $\Delta XYZ$ sea un mínimo.

Dar dos interpretaciones físicas diferentes.

**4.** Generalizar el ejemplo 3.

**5.** Criticar la solución del ejemplo 1. ¿Se aplicará a todos los casos?

**6.** Criticar la solución del ejemplo 3. ¿Se aplicará a todos los casos?

**7.** Dar una solución rigurosa del ejemplo 3 para triángulos acutángolos. (Variación parcial, ejemplos 1 y 5.)

**8.** Criticar las soluciones de la sección 1 (4) y sección 2 (2) para el problema del centro de tráfico. ¿Se aplicarán éstas a todos los casos?

**9.** *Centro de tráfico con cuatro puntos en el espacio.* Dado un tetraedro con vértices en los puntos $A$, $B$, $C$ y $D$. Asumir que hay un punto $X$ dentro del tetraedro tal que la suma de sus distancias desde los cuatro vértices

$$AX + BX + CX + DX$$

sea un mínimo. Mostrar que los ángulos $< AXB$ y $< CXD$ son iguales y tienen como bisectriz la misma línea recta; señalar otros pares de ángulos relacionados de modo semejante. [¿Conocemos un problema relacionado? ¿Un problema análogo? ¿Podría utilizarse su resultado o el método de su solución?]

**10.** *Centro de tráfico con cuatro puntos en el plano.* Considerar el caso extremo del ejemplo 9, cuyos puntos $A$, $B$, $C$ y $D$ están en el mismo

plano y son los cuatro vértices de un cuadrilátero convexo $ABCD$. ¿El enunciado del ejemplo 9 continuará siendo válido en este caso extremo?

**11.** *Red de tráfico con cuatro puntos.* Sean $A$, $B$, $C$ y $D$ cuatro puntos fijos, y $X$ e $Y$ dos puntos variables en un plano. Si el mínimo de la suma de las cinco distancias $AX + BX + XY + YC + YD$ es alcanzado de modo que los seis puntos $A$, $B$, $C$, $D$, $X$ e $Y$ son distintos, las tres líneas $XA$, $XB$ y $XY$ están igualmente inclinadas unas a otras y lo mismo ocurre con las tres líneas $YC$, $YD$ e $YX$.

**12.** *Desdoblar y estirar.* Hay todavía otra interpretación útil de la figura 9.3. Dibujar $l$, $A^*X$ y $XB$ sobre una hoja de papel transparente, después doblar la hoja a lo largo de la línea $l$: obtenemos así la figura 9.1 (con $A^*$ en lugar de $A$). Imaginemos la figura 9.1 originalmente dibujada bajo esta sofisticada manera una hoja transparente doblada. Al objeto de encontrar la posición de $X$ que hace de $AX + XB$ un mínimo, desdoblar la hoja, dibujar una línea recta desde $A$ (o más bien $A^*$ en la figura 9.3) a $B$ y luego volver a doblar la hoja.

**13.** *Billares.* Sobre una mesa rectangular de billar hay una bola en el punto $P$. Se quiere conducir la bola en una dirección tal que tras cuatro rechazos sucesivos sobre los cuatro lados del rectángulo la bola retorne a su posición original $P$. [Figura 9.14.]

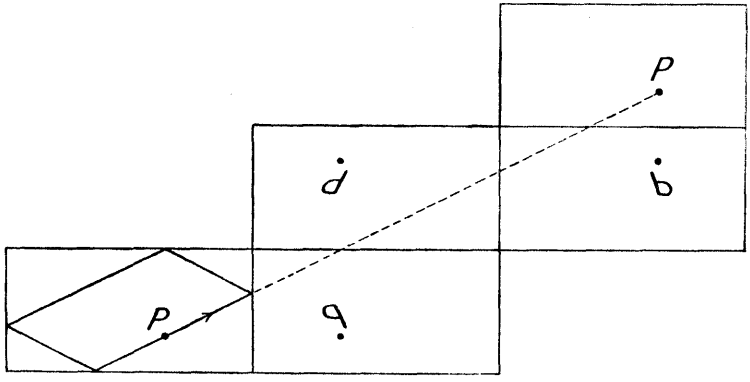



Fig. 9.14. La mesa de billar.

**14.** *Exploración geofísica.* En el punto $E$ de la superficie horizontal de la Tierra ocurre una explosión. El sonido de esta explosión es propagado en el interior de la Tierra y rechazado (reflejado) por una capa plana oblicua $OR$ que forma el ángulo $\alpha$ con la superficie de la

Tierra. El sonido que viene de $E$ puede llegar a un puesto de escuchas $L$ en otro punto de la superficie terrestre de $n$ modos diferentes. (Uno de estos $n$ caminos es dibujado en la figura 9.15 por el método del ejemplo 12.) Siendo dado $n$ (observado por un adecuado aparato), señalar los límites entre los que $\alpha$ se forma.

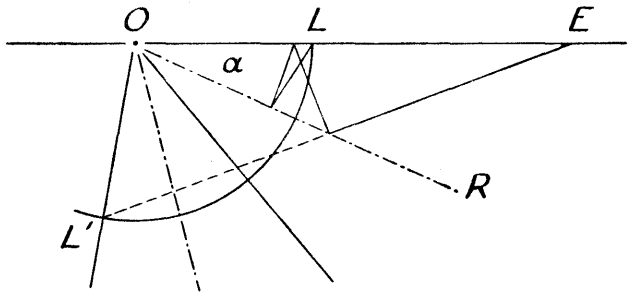



Fig. 9.15. Explosión subterránea.

**15.** Dados en el espacio una línea recta $l$ y dos puntos, $A$ y $B$, que no están en $l$, encontrar un punto $X$ en esta línea tal que la suma de sus distancias desde los dos puntos dados $AX + XB$ sea un mínimo.

¿Conocemos un problema relacionado? ¿Un problema más especial? ¿Podríamos utilizar su resultado o el método de su solución?

**16.** Resolver el ejemplo 15 usando una superficie de nivel tangente.

**17.** Resover el ejemplo 15 doblando un papel.

**18.** Resolver el ejemplo 15 por interpretación mecánica. ¿Está la solución de acuerdo con los ejemplos 16 y 17?

**19.** Dadas tres líneas rectas oblicuas en el espacio, $a$, $b$ y $c$, mostrar que el triángulo con un vértice en cada línea dada y un perímetro mínimo tiene la siguiente propiedad: la línea que une su vértice en la línea $a$ al centro de su círculo inscrito es perpendicular a $a$.

**20.** Considerar el caso particular del ejemplo 19 en que las tres líneas oblicuas son tres aristas de un cubo. ¿Dónde están los vértices del triángulo propuesto? ¿Dónde está el centro de su círculo inscrito? ¿Cuál es su perímetro si el volumen del cubo es $8a^3$?

**21.** Dadas tres líneas rectas oblicuas en el espacio, $a$, $b$ y $c$, sabemos que $X$ varía a lo largo de $a$, $Y$ a lo largo de $b$, $Z$ a lo largo de $c$, y $T$ libremente en el espacio. Encontrar el mínimo de $XT + YT + ZT$.

**22.** Especializar el ejemplo 21 como el ejemplo 20 especializó al 19.

**23.** *Las líneas más cortas en una superficie poliédrica.* Las paredes límites de una habitación rectangular son cuadradas; la habitación tiene 20 pies de larga, ocho de ancha y ocho de alta. Una araña está en una de las paredes finales a siete pies del suelo y a mitad de camino entre las paredes laterales. La araña percibe una mosca sobre la pared opuesta un pie sobre el suelo y también a mitad de distancia entre las paredes laterales. Mostrar que la araña tiene que andar menos de 28 pies a lo largo de las paredes, el techo o el suelo, para alcanzar la mosca. [Ejemplo 17.]

**24.** *Las líneas más cortas (geodésicas) en una superficie curvada.* Consideramos una especie curvada como el límite de un poliedro. Cuando el poliedro se aproxima a la superficie curvada el número de sus caras tiende a $\infty$, la diagonal más larga de cualquier cara tiende a 0 y las caras tienden a ser tangentes a la superficie.

Sobre una superficie poliédrica la línea más corta entre dos puntos es un polígono. Puede ser un polígono plano aquel cuyos puntos yacen en un plano, o puede ser un polígono oblicuo aquel cuyos puntos no están contenidos en un plano. (Ambos casos pueden ser ilustrados por la solución del ejemplo 23, el primer caso por (1), el segundo caso por (2) y (3)).

Sobre una superficie curvada la línea más corta recibe el nombre de «geodésica» porque la línea más corta juega un papel en geodesia, la ciencia que estudia la superficie terrestre. La geodésica puede ser una curva plana, contenida por completo en un plano, o puede ser una «curva espacial» (curva «tortuosa») cuyos puntos no pueden estar contenidos en un plano. En todo caso, una geodésica debe tener alguna intrínseca relación geométrica con la superficie en la que ésta es la línea más corta. ¿Qué relación es ésta?

(1) Consideremos la línea poligonal $ABC \ldots L$. Aun si $ABC \ldots L$ es un polígono oblicuo, dos segmentos consecutivos de éste, como $HI$ e $IJ$, yacen en el mismo plano. Si $ABC \ldots L$ es la línea más corta sobre una superficie poliédrica entre sus puntos finales $A$ y $L$, cada uno de sus vértices $B, C, D, \ldots H, I, J, \ldots K$ están sobre una arista del poliedro. El plano que contiene los segmentos $HI$ a $IJ$ contienen también la bisectriz de $< HIJ$, y esta bisectriz es perpendicular a la arista del poliedro que pasa por $I$; ver el ejemplo 16 o el 18.

Consideremos una curva. Aun si la curva es tortuosa, un arco infinitesimal (muy corto) puede ser considerado como un arco plano (casi plano). El plano del arco infinitesimal es el *plano osculador* en su punto medio. El plano osculador es análogo al plano en que se tocan dos segmentos sucesivos de un polígono oblicuo. Si la curva es geodésica,

es decir, la línea más corta sobre una superficie, la analogía sugiere que el *plano osculador de una geodésica en un punto pasa por la normal a la superficie en este punto.*

(2) Una geodésica puede ser interpretada físicamente como una cinta de goma estirada a lo largo de una superficie suave (sin fricciones). Examinemos el equilibrio de una pequeña porción de la cinta de goma. Las fuerzas sobre esta porción son dos tensiones iguales que actúan tangencialmente a los dos puntos finales del corto arco y la reacción de la superficie sin fricción actúa sobre él normalmente. La reacción de la superficie, compuesta por una fuerza resultante, y las dos tensiones en los puntos finales están en equilibrio. Por tanto, estas tres fuerzas son paralelas al mismo plano. Sin embargo, dos «tangentes vecinas» determinan el plano osculador que, en consecuencia, contiene la normal a la superficie.

(3) Cada arco de una geodésica es una geodésica. En efecto, si en una curva hay una porción que no es la más corta entre sus puntos finales y, por tanto, puede ser remplazado por un arco más corto entre los mismos puntos, la curva total no puede ser la línea más corta. De aquí que sea natural que una geodésica tenga alguna propiedad distintiva en cada uno de sus puntos. La propiedad sugerida por dos consideraciones heurísticas muy diferentes, (1) y (2), es una propiedad de esta clase.

(4) Busquemos ejemplos para comprobar el resultado obtenido heurísticamente. ¿Cuáles son las líneas más cortas en una esfera? ¿Tienen éstas la propiedad sugerida? ¿Otras líneas sobre la superficie esférica tienen la misma propiedad?

**25.** Un punto material se mueve sin fricción sobre una superficie suave y rígida. Ninguna fuerza exterior (tal como la gravitación) actúa sobre el punto (excepto, por supuesto, la reacción de la superficie). Dar razones por las que se espera que el punto describa una geodésica.

**26.** *Una construcción con papel doblado.* Encontrar un polígono inscrito en un círculo cuando los lados están dados en magnitud y en sucesión.

$a_1, a_2, a_3, \ldots, a_n$ denotan las longitudes dadas. El lado de longitud $a_1$ es seguido por el de longitud $a_2$, éste por el de longitud $a_3$, y así sucesivamente; el lado de longitud $a_n$ es seguido por el de longitud $a_1$. Se comprende que cualquiera de las longitudes $a_1, a_2, \ldots, a_n$ es menor que la suma de las longitudes que quedan $n - 1$.

Hay una bonita solución con papel doblado. Dibujar $a_1, a_2, \ldots, a_n$ sobre cartulinas como cuerdas sucesivas en un círculo suficientemente

largo de modo que dos cuerdas consecutivas tengan un punto común. Dibujar los radios de estos puntos finales al centro de círculo. Recortar el polígono limitado por $n$ cuerdas y los dos radios extremos, doblar la cartulina a lo largo de los otros $n-1$ radios, y pegar juntos los dos radios por los que fue cortada la cartulina. Obtenemos así una superficie poliédrica abierta; ésta consta de $n$ rígidos triángulos isósceles, es limitado por $n$ aristas libres de longitudes $a_1, a_2, \ldots, a_n$, respectivamente, y tiene $n$ ángulos diedros que aún pueden ser cambiados. (Suponemos que $n > 3$.)

¿Qué podemos hacer con esta superficie poliédrica para resolver el problema propuesto?

**27.** *La suerte está echada.* La masa en el interior de un pesado poliedro convexo no necesita ser uniformemente repartida. En efecto, podemos imaginar una distribución heterogénea y conveniente de la masa cuyo centro de gravedad coincide con un punto interior del poliedro arbitrariamente asignado. Puesto sobre el suelo, el poliedro descansará en una de sus caras. Esto nos proporciona un argumento mecánico para la siguiente proposición geométrica.

Dado un poliedro convexo $P$ y un punto $C$ en el interior de $P$, podemos encontrar una cara $F$ de $P$ con las siguientes propiedades: el pie de la perpendicular dibujada desde $C$ al plano de $F$ es un punto *interior* de $F$.

Encontrar una prueba geométrica de esta proposición. (Observemos que la cara $F$ puede, pero no necesita, ser únicamente determinada por la propiedad establecida.)

**28.** *El Diluvio.* Has tres clases de puntos observables en un mapa de contornos: picos, pasos (puntos de nivel en un plano horizontal tangente) y «profundidades». (En la figura 8.7, $P$ es un pico y $S$ un paso.) Una «profundidad» es el punto más bajo en el fondo de un valle desde el cual el agua no encuentra salida. Una profundidad es un pico «invertido»: en el mapa de contornos miramos cualquier línea de elevación $h$ como si tuviese $-h$ de elevación. Entonces el mapa está «invertido»; parece el mapa de un paisaje bajo el mar, los picos se convierten en profundidades, las profundidades en picos, pero los pasos permanecen iguales. Hay una observable conexión entre estas tres clases de puntos.

*Supongamos que hay $P$ picos, $D$ profundidades y $S$ pasos en una isla. Luego,*

$$P + D = S + 1.$$

Al objeto de derivar este teorema intuitivamente imaginemos que

una lluvia persistente aumenta el agua de la isla hasta que toda ella queda sumergida. Podemos admitir que todos los picos $P$ son igualmente altos y que todas las profundidades están sobre, o bajo, el nivel del mar. De hecho, podemos imaginar los picos levantados y las profundidades deprimidas sin cambiar ningún número. Cuando empezó a llover el agua se almacenaba en las profundidades; tenemos

$$D + 1 \text{ capas de agua y } 1 \text{ isla}$$

al principio. Un poco antes de que la isla fuese sumergida, sobre el agua sólo se veían los picos, y, así, tenemos,

$$1 \text{ capa de agua y } P \text{ islas}$$

al fin. ¿Cómo ha ocurrido la transición?

Imaginemos que, en cualquier momento, las distintas capas de agua están a la misma elevación. Si no hay paso precisamente en esta elevación, el agua puede levantarse un poco más sin que cambie ni el número de capas de agua ni el número de islas. No obstante, cuando el agua alcanza un nuevo paso, el mínimo levantamiento de su nivel unirá dos capas de agua separadas o trozos de tierra. Por tanto, cada paso de agua o disminuye el número de capas de agua en una unidad o aumenta en una unidad el número de islas. Mirando el cambio total, obtenemos

$$(D + 1 - 1) + (P - 1) = S,$$

que es el teorema deseado.

(*a*) Supongamos ahora que hay $P$ picos, $D$ profundidades y $S$ pasos sobre todo el globo (algunos bajo el agua) y mostrar que

$$P + D = S + 2.$$

(*b*) La última relación nos recuerda el teorema de Euler (véanse secciones 3.1-3.7 y ejemplos 3.1-3.9). ¿Podríamos usar el teorema de Euler para construir una prueba geométrica del resultado obtenido por un argumento intuitivo? [Las figuras 9.16 y 9.17 muestran piezas importantes de un mapa más completo, en el que no sólo son indicadas algunas líneas de nivel, sino también algunas «líneas de escarpados descendentes» que son perpendiculares a las líneas de nivel. Estas dos clases de líneas subdividen la superficie del globo en triángulos y cuadriláteros. Cf. el ejemplo 3.2.]

(*c*) ¿Hay algunas observaciones sobre el método?

**29.** *No tan profundo como un pozo.* Al objeto de encontrar $d$, lo profundo de un pozo, podemos dejar caer una piedra dentro y medir el tiempo $t$ entre el momento de soltar la piedra y el momento en que la oímos chocar contra el agua.

(*a*) Dado $g$, la aceleración gravitatoria, y $c$, la velocidad del sonido,

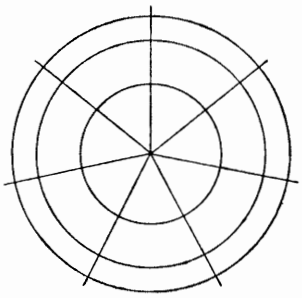


Fig. 9.16. Alrededores de un pico.

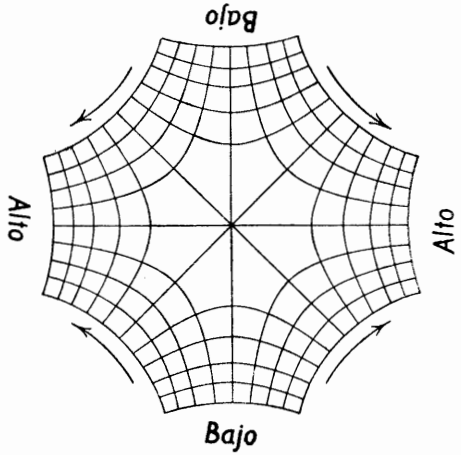



Fig. 9.17. Alrededores de un paso.

expresar $d$ en términos de $g$, $c$ y $t$. (Desdeñar la resistencia del aire.)

(*b*) Si el pozo no es demasiado profundo, la velocidad final de la piedra será una pequeña fracción de la velocidad del sonido, y así podemos esperar que la mayor parte de la medida del tiempo $t$ es tomada por la caída de la piedra. De aquí podríamos esperar que

$$d = gt^2/2 - \text{co rrección}$$

donde la corrección es relativamente pequeña cuando $t$ es pequeño.

Al objeto de examinar esta intuición extendemos la expresión obtenida como contestación a (*a*) en valores de $t$ y retener los dos primeros términos, que no desaparecen.

¿Qué podemos considerar como característico en este ejemplo?

**30.** *Un caso extremo útil.* Una elipse que gira sobre su eje mayor describe un esferoide alargado, o elipsoide de revolución con forma de huevo. Los focos de la elipse rotativa no giran: están sobre el eje de revolución y son llamados también focos del esferoide oblongo. Podríamos hacer un espejo elíptico convirtiendo el lado cóncavo interior de la superficie con un metal pulido; toda luz que viene de un foco es reflejada en el otro foco por un espejo elíptico tal; véase la sección 1 (3). Los espejos

elípticos son muy rara vez usados en la práctica, pero hay un caso límite que es muy importante en astronomía. ¿Qué sucede si uno de los focos del elipsoide está fijo y el otro tiende a infinito?

**31.** Resolver la ecuación diferencial de la braquistocrona encontrada en la sección 4.

**32.** *El cálculo de variaciones* se refiere a problemas de máximos y mínimos de cantidades que dependen de la forma y el tamaño de una curva variable. Tal es el problema de la braquistocrona resuelto en la sección 4 por interpretación óptica. El problema de las geodésicas o líneas más cortas sobre una superficie curva, discutidos en el ejemplo 24, también pertenece al cálculo de variaciones, y el «problema isoperimétrico», que será tratado en el próximo capítulo, pertenece también al mismo. Consideraciones físicas, que pueden resolver varios problemas de máximos y mínimos, como hemos visto, pueden también resolver algunos problemas del cálculo de variaciones. Esbocemos un ejemplo.

*Encontrar la curva con longitud dada y puntos finales dados que tiene un centro de gravedad de elevación mínima.* Se asume que la densidad de materia pesada es constante a lo largo de la curva que podemos mirar como una cuerda o cadena uniforme. Cuando el centro de gravedad de la cadena alcanza su posición más baja posible, la cadena está en *equilibrio*. Ahora podemos averiguar el equilibrio de la cadena examinando las fuerzas que actúan sobre ella, su peso y su tensión. Esta investigación conduce a una ecuación diferencial que determina la curva deseada, *la catenaria*. No vamos a entrar en detalles. Sólo deseamos poner de relieve que la solución esbozada tiene la misma idea básica que las soluciones mecánicas consideradas en la sección 2.

**33.** *Del equilibrio de las secciones transversales al equilibrio de los sólidos.* Arquímedes no estableció explícitamente el principio general de su método, pero lo aplicó a varios ejemplos, cálculo de volúmenes, áreas y centros de gravedad, y la variedad de estas aplicaciones hace el principio perfectamente claro. Apliquemos la variante del método de Arquímedes, que ha sido presentada en la sección 5, a algunos de sus ejemplos.

Probar la proposición 7 del *Método:* Cualquier segmento de la esfera tiene con respecto al cono de la misma base y altura la misma razón que la suma del radio de la esfera y la altura del segmento complementario tiene con respecto a la altura del segmento complementario.

**34.** Probar la proposición 6 del *Método:* El centro de gravedad de un hemisferio está en su eje y lo divide de modo que la posición adya-

cente al vértice del hemisferio tiene con la porción restante la relación de 5 a 3.

**35.** Probar la proposición 9 del *Método:* El centro de gravedad de cualquier segmento de la esfera está en su eje y lo divide de modo que la porción adyacente al vértice tiene con respecto a la parte restante la misma relación que la suma del eje del segmento y cuatro veces el eje del segmento complementario tiene con respecto a la suma del eje del segmento y el doble del eje del segmento complementario.

**36.** Probar la proposición 4 del *Método:* Cualquier segmento de un paraboloide de revolución cortado por un plano en ángulos rectos al eje tiene la relación de 3 a 2 con respecto al cono de la misma base y altura que el segmento.

**37.** Probar la proposición 5 del *Método:* El centro de gravedad de un segmento de un paraboloide de revolución, cortado por un plano en ángulos rectos al eje, está sobre éste y lo divide de modo que la parte adyacente al vértice es el doble de la parte restante.

**38.** *El Método de Arquímedes, retrospectivamente.* Lo que Arquímedes tenía en la mente cuando descubrió su método nunca lo sabremos y sólo vagamente podemos intuirlo. Sin embargo, podemos proponer una clara y corta lista de las reglas matemáticas (bien conocidas hoy, pero no formuladas aún en tiempo de Arquímedes) que necesitamos para resolver, con métodos contemporáneos, los problemas que Arquímedes resolvió con su método. Necesitamos:

(1) Dos reglas generales de cálculo integral:

$$\int cf(x)dx = c\int f(x)dx, \quad \int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx;$$

$c$ es una constante, $f(x)$ y $g(x)$ son funciones.

(2) El valor de cuatro integrales:

$$\int x^n dx = x^n + 1(n + 1) \qquad \text{para } n = 0, 1, 2, 3.$$

(3) La interpretación geométrica de dos integrales:

$$\int Q(x)dx, \qquad \int xQ(x)dx.$$

$Q(x)$ denota una longitud en geometría plana y un área en geometría sólida; en ambos casos denota la sección transversal variable de una figura determinada por un plano perpendicular al eje de las $x$. La primera integral expresa un área o volumen; la segunda integral,

el momento de un área o volumen uniformemente lleno, según consideremos un problema de geometría plana o sólida.

Arquímedes no formuló estas reglas, aunque sin duda él las poseyó de una u otra forma. El se abstuvo de formular en términos generales el proceso subyacente, el paso de la sección variable al área o el volumen, de lo integrante a lo integral, como diríamos hoy. Por el contrario, describió este proceso en casos particulares, lo aplicó en una gran variedad de casos, sin duda lo conocía íntimamente, pero siempre lo consideró como meramente heurístico y pensó que esto era una buena razón para abstenerse de formularlo en general.

Citar sencillos valores geométricos de los que se desprendan intuitivamente los valores de las cuatro integrales señaladas en (2)*,

---

* Para otras observaciones sobre el descubrimiento de Arquímedes cf. B. L. VAN DER WAERDEN, *Elemente der Mathematik*, vol. 8, 1953, págs. 121-129, y vol. 9, 1954, páginas 1-9.

# El problema isoperimétrico

*El círculo es la primera figura, la más sencilla y la más perfecta.*—PROCLUS *.

*Lo cerchio è perfettissima figura.*—DANTE **.

## 1. Las razones inductivas de Descartes

En la inacabada obra de Descartes *Regulae ad Directionem Ingenii* (o *Reglas para la dirección de la mente*, que, por cierto, debe ser considerada como una de las obras clásicas sobre lógica del descubrimiento) encontramos este curioso pasaje ***: «Al objeto de mostrar por enumeración que el perímetro de un círculo es menor que el de cualquiera otra figura del mismo área no necesitamos un examen completo de todas las figuras, sino que basta probar esto para unas pocas figuras particulares, de las que podemos concluir lo mismo, por inducción, para todas las demás figuras».

Para comprender el significado del pasaje, realicemos, en realidad, lo que Descartes sugiere. Comparemos el círculo a unas pocas otras figuras, triángulos, rectángulos y sectores circulares. Tomemos dos triángulos: el equilátero y el rectángulo isósceles (con ángulos de 60°, 60°, 60° y 90°, 45°; 45°; respectivamente). La forma de un rectángulo es caracterizada por la razón de su ancho a su alto; elegimos las razones 1 : 1 (cuadrado), 2 : 1, 3 : 1 y 3 : 2. La forma de un sector del círculo está determinada por el ángulo del centro; elegimos los ángulos 180°, 90° y 60° (semicírculo, cuadrante y sextante). Asumimos que todas estas figuras tienen el mismo área, digamos, una pulgada cuadrada. Luego calculamos la longitud del perímetro de cada figura en pulgadas. Los números obtenidos están recogidos en la siguiente tabla; las figuras están ordenadas de modo que los perímetros aumentan de arriba abajo.

---

* Comentario al primer libro de los *Elementos*, de EUCLIDES; definiciones XV y XVI.

** *Convivio*, II, XIII, 26.

*** *Oeuvres de Descartes*, editada por Adam y Tannery, vol. 10, 1908, pág. 390. El pasaje es ligeramente alterado; la propiedad del círculo en consideración es establecida aquí en forma diferente.

TABLA I

PERIMETROS DE LAS FIGURAS DE IGUAL AREA

| | |
|---|---|
| Círculo | 3,55 |
| Cuadrado | 4,00 |
| Cuadrante | 4,03 |
| Rectángulo 3 : 2 | 4,08 |
| Semicírculo | 4,10 |
| Sextante | 4,21 |
| Rectángulo 2 : 1 | 4,24 |
| Triángulo equilátero | 4,56 |
| Rectángulo 3 : 1 | 4,64 |
| Triángulo rectángulo isósceles | 4,84 |

De todas las figuras enumeradas que tienen el mismo área, el círculo, que está en primer lugar, tiene el perímetro más corto. ¿Podemos concluir de aquí por inducción, como Descartes parece sugerir, que el círculo tiene el perímetro más corto no sólo entre las figuras enumeradas, sino entre todas las figuras posibles? De ningún modo. Pero no podemos negar que nuestra lista relativamente corta sugiere con mucha fuerza el teorema general; con tanta fuerza que, si añadiésemos una o dos figuras más a la lista, la sugestión no sería mucho más fuerte.

Por mi parte me siento inclinado a creer que Descartes, al escribir el pasaje citado, pensó en este último y sutil punto. El quiso decir, seguramente, que la prolongación de la lista no influiría mucho en nuestra creencia.

## 2. RAZONES LATENTES

«De todas las figuras planas de igual área el círculo es el de perímetro mínimo.» Llamemos a esta proposición, apoyada por la tabla I, el *teorema isoperimétrico**. La tabla I, construida de acuerdo con la sugerencia de Descartes, nos da un argumento inductivo claramente convincente en favor del teorema isoperimétrico. Sin embargo, ¿por qué el teorema no nos convence?

Imaginemos una situación de algún modo semejante. Elegimos diez árboles de diez familiares diferentes. Medimos el peso específico de la madera de cada árbol y separamos los árboles cuya madera tiene menos peso específico. ¿Sería razonable creer meramente sobre la base

* Una explicación del nombre y formas equivalentes será dada más adelante (sección 8).

de estas observaciones que la clase de madera que tiene peso más ligero entre las diez examinadas serán también las de menos peso entre las maderas de todos los árboles existentes? Creer esto no sería razonable, sino tonto.

¿Cuál es la diferencia con el caso del círculo? Estamos *prejuiciados* en favor del círculo. El círculo es la figura más perfecta; realmente, creemos que, entre todas sus perfecciones, el círculo tiene el perímetro más corto para un área dada. El argumento inductivo sugerido por Descartes parece tan convincente porque corrobora una conjetura plausible desde el comienzo.

«El círculo es la figura más perfecta» es una frase tradicional. La encontramos en los escritos de Dante (1265-1321), de Proclo (410-485) y aun de otros escritores anteriores. El significado de la frase no es claro, pero puede tener detrás algo más que la mera tradición.

## 3. Razones físicas

«De todos los sólidos de igual volumen la esfera tiene la superficie mínima.» Llamaremos esta proposición el «teorema isoperimétrico en el espacio».

Estamos inclinados a creer en el teorema isoperimétrico en el espacio, como en el plano, sin ninguna demostración matemática. Estamos prejuiciados en favor de la esfera quizá por lo mismo que lo estamos en favor del círculo. De hecho, la misma naturaleza parece estar prejuiciada en favor de la esfera. Las gotas de lluvia, las pompas de jabón, el Sol, la Luna, nuestro globo, los planetas, son esféricos, o casi esféricos. Con poco conocimiento de la física de tensión de superficies podemos aprender el teorema isoperimétrico en una pompa de jabón.

Aun siendo ignorantes de física, podemos llegar al teorema isoperimétrico por consideraciones bastante primitivas. Podemos aprenderlo también de un gato. Seguramente todos hemos visto un gato cuando se prepara para dormir en una noche fría: estira sus patas, se enrosca y, en definitiva, hace su cuerpo tan esférico como es posible. El lo hace así, desde luego, para guardar calor, para hacer mínimo el calor que escapa por su superficie corporal. El gato, que no desea disminuir su volumen, intenta disminuir su superficie. Resuelve por su cuenta el problema de un cuerpo con volumen dado y superficie mínima haciéndose a sí mismo tan esférico como sea posible. Parece que tiene algún conocimiento del teorema isoperimétrico.

La física que subyace a esta consideración es extremadamente tosca *. Aun así, la consideración es convincente y válida, como una especie de apoyo provisional para el teorema isoperimétrico. Las razones elusivas que hablan en favor de la esfera o el círculo, apuntadas arriba (véase sección 2), están condensadas. ¿Son éstas razones de analogía física?

## 4. Las razones inductivas de lord Rayleigh

Un poco más de doscientos años después de la muerte de Descartes, el físico lord Rayleigh investigó los tonos de las membranas. El pergamino estirado sobre un tambor es una «membrana» (o, más bien, una razonable aproximación a la idea matemática de membrana) siempre que sea hecho muy cuidadosamente y estirada uniformemente a lo largo de él. Los tambores son normalmente de forma circular, pero, después de todo, no hay razones para que no podamos hacerlos con forma elíptica, o poligonal, o cualquier otra. Un tambor de cualquier forma puede producir diferentes tonos, de los cuales el más profundo, llamado tono principal, es el más fuerte. Lord Rayleigh comparó los tonos principales de membranas de diferentes formas, pero de igual área y sujetos a las mismas condiciones físicas. El construyó la siguiente tabla II, muy parecida a la tabla I de la sección 1. Esta tabla II contiene las mismas formas que la tabla I, aunque colocadas en un orden algo diferente, y da por cada forma el diapasón (frecuencia) del tono principal **.

Tabla II

PRINCIPALES FRECUENCIAS DE MEMBRANAS DE AREA IGUAL

| | |
|---|---|
| Círculo | 4,261 |
| Cuadrado | 4,443 |
| Cuadrante | 4,551 |
| Sextante | 4,616 |
| Rectángulo 3 : 2 | 4,624 |
| Triángulo equilátero | 4,774 |
| Semicírculo | 4,803 |
| Rectángulo 2 : 1 | 4,967 |
| Triángulo rectángulo isósceles | 4,967 |
| Rectángulo 3 : 1 | 5,736 |

* Un gato mejor aconsejado no haría de la superficie de su cuerpo un mínimo, sino su conductibilidad térmica o, lo que es lo mismo, su capacidad electrostática. Aún, mediante un teorema de Poincaré, este diferente problema de mínimos tiene la misma solución: la esfera. Véase G. Polya, *American Mathematical Monthly*, v. 54, 1947, págs. 201-206.

** Lord Rayleigh, *The Theory of Sound*, 2.ª ed., vol. 1, pág. 345.

De las diez membranas alineadas, todas del mismo área, la membrana circular, puesta en primer lugar, es la de tono principal más profundo. ¿Podemos concluir de aquí por inducción que el círculo tiene el tono principal más bajo de *todas* las formas?

Por supuesto que no; la inducción nunca es conclusiva. Sin embargo, la sugestión es muy fuerte, aún más fuerte que en el caso anterior. Sabemos (y lord Rayleigh y sus contemporáneos también lo sabían) que de todas las figuras con un área dada el círculo tiene el perímetro mínimo, y que este teorema puede ser demostrado matemáticamente. Con esta propiedad geométrica mínima del círculo en la mente, nos sentimos inclinados también a creer que el círculo tiene la propiedad física mínima sugerida por la tabla II. Nuestro juicio está influido por la analogía, y la analogía tiene una profunda influencia.

La comparación de la tabla I y de la II es altamente instructiva; nos proporciona otras varias sugestiones que no intentaremos discutir ahora.

## 5. Derivando consecuencias

Hemos examinado varios fundamentos en favor del teorema isoperimétrico, que son insuficientes para probarlo, pero suficientes para hacer de él una conjetura razonable. Un físico, examinando una conjetura de su ciencia, deriva consecuencias de ella. Estas consecuencias pueden o no estar de acuerdo con los hechos y el físico ingenia experimentos para averiguar en qué caso se encuentran. Un matemático, examinando una conjetura de su ciencia, puede seguir un curso semejante. El deriva consecuencias de su conjetura. Estas consecuencias pueden o no ser ciertas y el matemático trata de averiguar cuál es el caso.

Seguimos este curso al examinar el teorema isoperimétrico que hemos establecido ahora en la siguiente forma: *De todas las figuras planas de igual perímetro el círculo tiene el área mínima.* Este teorema difiere del que dimos arriba (sección 2) y no sólo verbalmente. Sin embargo, podemos mostrar que ambos enunciados son equivalentes. Por ahora posponemos la prueba (sección 8) y corremos a examinar las consecuencias.

(1) Dido, la fugitiva hija del rey tirio, llegó después de muchas aventuras a las costas de Africa, convirtiéndose en la fundadora de Cartago y su primera reina legendaria. Dido empezó por comprar a los nativos un trozo de tierra sobre el litoral «no más largo que lo que

una piel de toro podía rodear». Dido cortó la piel en finas y estrechas tiras, con las que hizo una cuerda muy larga. Y entonces se le presentó a Dido un problema geométrico: ¿Qué forma de tierra acotaría con su cuerda, de longitud dada, para obtener un área máxima?

En el interior del continente la contestación sería, por supuesto, un círculo, pero en el litoral la cuestión es diferente. Resolvámoslo asumiendo que el litoral es una línea recta. En la figura 10.1 el arco $XYZ$ tiene una longitud dada. Se nos pide hacer un máximo del área comprendida entre este arco y la línea recta $XZ$ (que está sobre una línea recta dada e infinita, pero que puede ser alargada o acortada a placer).

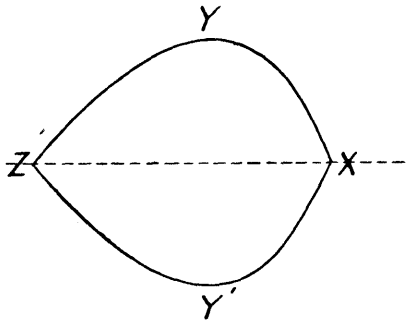



Fig. 10.1 El problema de Dido.

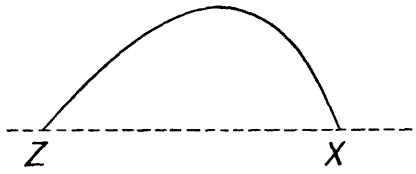



Fig. 10.2. Solución por la imagen de espejo.

Para resolver este problema consideremos la línea recta dada e infinita (el litoral) como un espejo; véase la figura 10.2. La línea $XYZ$ y su imagen en el espejo $XY'Z$ forman en conjunto una *cerrada* curva $XYZY'$ de longitud dada que rodea un área que mide exactamente el doble que la que deseamos maximizar. Este área es un máximo cuando la curva cerrada es un círculo cuya línea recta dada e infinita (el litoral) es un eje de simetría. Por tanto, la solución del problema de Dido es un semicírculo con centro en el litoral.

(2) Jacob Steiner derivó un grupo de interesantes consecuencias del teorema isoperimétrico. Vamos a discutir uno de sus argumentos, que es especialmente sorprendente. Inscribamos en un círculo dado un polígono (figura 10.3). Consideremos rígidos (cortados sobre cartulina) los segmentos del círculo (sombreados en la figura 10.3) separados por los lados del polígono inscrito. Imaginemos estos rígidos segmentos del círculo conectados por flexibles junturas en los vértices del polígono inscrito. Deformar este articulado sistema cambiando los ángulos en las junturas. Tras la deformación obtenemos una nueva curva que no es un círculo (ver la figura 10.4), sino que consiste en

sucesivos arcos circulares y tiene el mismo perímetro que el círculo dado. Por tanto, según el teorema isoperimétrico, el área de la nueva curva debe ser menor que el área del círculo dado. Sin embargo, los segmentos circulares son rígidos (de cartulina), sus áreas no cambian, y de esta forma el polígono deformado disminuye el área: *El área de*

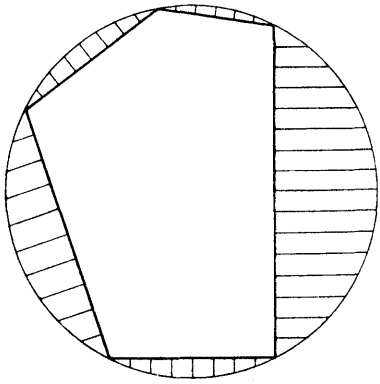


Fig. 10.3. Un polígono escrito.

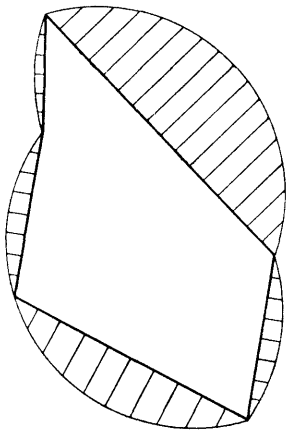


Fig. 10.4. Junturas flexibles y segmentos de cartulina.

*un polígono inscrito en un círculo es más grande que el área de cualquier otro polígono con los mismos lados.* (Los lados son iguales en longitud y en orden de sucesión.)

Esta consecuencia es bella, pero todavía no ha sido probada, así como no hemos probado el teorema isoperimétrico mismo.

(3) Combinemos el problema de Dido con el método de Steiner. Inscribamos en un semicírculo dado una línea poligonal; ver la figu-

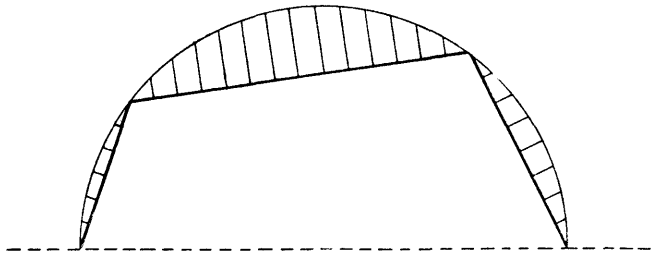


Fig. 10.5. Dido y Steiner.

ra 10.5. Consideremos los segmentos cortados por tramos de la línea poligonal a partir del semicírculo (sombreado en la figura 10.5) como totalmente rígidos (de cartulina). Coloquemos junturas flexibles en los

vértices de la línea poligonal, variemos los ángulos y llevemos los puntos finales a lo largo de la línea del diámetro, que consideraremos como dada. Obtenemos así una nueva curva (figura 10.6), que consta de arcos circulares de la misma longitud total que el semicírculo, pero que forman con la línea infinita dada menos área que la del semicírculo, por virtud del teorema que hemos discutido en (1). Pero los segmentos circulares son rígidos (de cartulina) y, por tanto, sólo el polígono deformado es responsable por la disminución del área. De aquí el teorema: *Los lados de un polígono, excepto uno, son dados en longitud y en sucesión. El área alcanza un máximo cuando el polígono es inscrito en un semicírculo cuyo diámetro es el lado originalmente no dado.*

## 6. Verificando consecuencias

Un físico, habiendo derivado varias consecuencias de su conjetura, busca una que pueda ser convenientemente comprobada por experimentos. Si los experimentos claramente contradicen una conjetura,

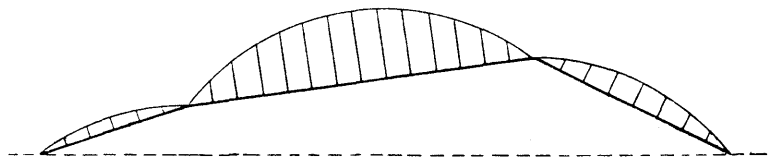


Fig. 10.6. Los segmentos son de cartulina.

esta misma queda refutada. Si los experimentos verifican sus consecuencias, la conjetura gana autoridad, adquiere más crédito. El matemático puede seguir un curso semejante. Este busca consecuencias accesibles de su conjetura que puedan probarla o refutarla. Una consecuencia refutada refuta la conjetura misma. Una consecuencia probada hace la conjetura más digna de crédito y establece un lazo a cuyo través la conjetura misma puede ser probada.

¿Qué pasa con nuestro caso? Hemos derivado varias consecuencias del teorema isoperimétrico: ¿cuál es la más accesible?

(1) Algunas de las consecuencias derivadas del teorema isoperimétrico en la sección anterior se refieren, en efecto, a problemas elementales de máximos. ¿Hay alguna consecuencia que podamos verificar? Examinemos los distintos casos indicados por las figuras 10.3-10.6. ¿Qué caso es el más sencillo? La complejidad de un polígono aumenta con el número de sus lados. Por tanto, el polígono más sencillo de todos es el

triángulo; naturalmente, preferimos el triángulo porque sabemos más sobre él. Ahora bien, el problema de las figuras 10.3 y 10.4 no tiene sentido para triángulos o podemos decir que es vacuo en el caso de un triángulo: un triángulo con dados lados es determinado, rígido. No hay transición para un triángulo como ese de la figura 10.3 a la 10.4. Sin embargo, la transición de la figura 10.5 a la 10.6 es perfectamente posible para triángulos. Esta puede ser la consecuencia más sencilla que hemos derivado del teorema isoperimétrico: vamos a examinarla.

El caso particular más sencillo del resultado derivado de la sección 5 (3) responde al siguiente problema: *Dados dos lados de un triángulo encontrar el máximo del área;* véase figura 10.7. La sección 5 (3) da esta respuesta: el área es un máximo cuando el triángulo está inscrito en un semicírculo cuyo diámetro es el lado no dado originalmente. Esto significa que el área es un máximo cuando los dos lados dados forman un ángulo recto, lo que es obvio (ejemplo 8.7).

Hemos logrado verificar una primera consecuencia del teorema isoperimétrico. Este éxito levanta nuestros ánimos. ¿Qué hay detrás del hecho que acabamos de verificar? ¿Podríamos generalizarlo? ¿Podríamos verificar alguna otra consecuencia?

(2) Al generalizar el problema discutido en (1), llegamos a lo siguiente: *Todos los lados sucesivos de un polígono son dados en magnitud, excepto uno. Encontrar el área máxima.*

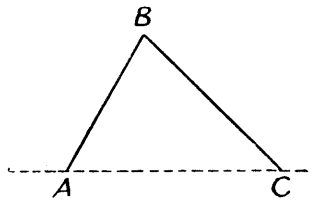



Fig. 10.7. El dedo con una articulación.

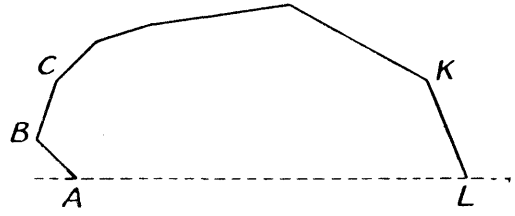



Fig. 10.8. El superdedo.

Introduzcamos una notación conveniente y dibujemos la figura 10.8. Las longitudes $AB$, $BC$, ..., $KL$ son dadas; la longitud $LA$ no está dada. Podemos imaginar la línea quebrada $ABC \ldots F \ldots KL$ como una especie de «superdedo»; los «huesos» $AB$, $BC$, ... $KL$ son de longitud invariable, los ángulos de las articulaciones $B$, $C$, ... $F \ldots K$ son variables. Se nos pide que hagamos del área $ABC \ldots KLA$ un máximo.

Como en algunos problemas que consideramos hace tiempo (secciones 8.4, 8.5), la dificultad característica parece ser que hay muchas

variables (los ángulos de $B$, $C$, ... $F$ ... y $K$). Sin embargo, acabamos de discutir en (1) el caso particular extremo del problema en que hay sólo un ángulo variable (sólo una articulación; figura 10.7). Es natural esperar que podamos utilizar este caso particular como un primer paso hacia la solución del problema general.

De hecho, tomemos el problema como casi resuelto. Imaginemos que hemos obtenido ya los valores deseados de todos los ángulos menos

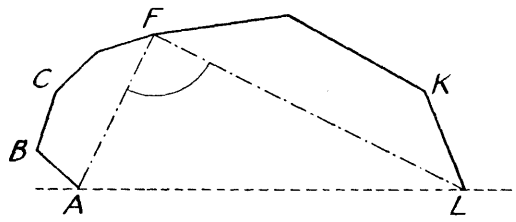



Fig. 10.9. Sólo una articulación es flexible.

uno. En la figura 10.9 consideramos el ángulo $F$ como variable, pero todos los demás, $B$, $C$, ... $K$ como fijos; las articulaciones $B$, $C$, ... $K$ son rígidas, sólo $F$ es flexible. Unamos $A$ y $L$ a $F$. Las longitudes $AF$ y $LF$ son invariables. El polígono total $ABC$ ... $F$ ... $KLA$ queda ahora descompuesto en tres partes, dos de las cuales son rígidas (de cartulina) y sólo la tercera es variable. Los polígonos $ABC$ ... $FA$ y $LK$ ... $FL$ son rígidos. El triángulo $AFL$ tiene dos lados dados, $FA$ y $FL$, y un ángulo variable en $F$. El área de este triángulo, y con él el área de todo el polígono $ABC$ ... $F$ ... $KLA$, alcanza un máximo cuando $< AFL$ es un ángulo recto, como hemos dicho hace un momento, en (1), al examinar la figura 10.7.

Este razonamiento se aplica también a las otras articulaciones, es decir, a los ángulos en $B$, $C$, ... y $K$ (figura 10.8), y así vemos que: *el área del polígono $ABC$ ... $KLA$ no puede ser un máximo a menos que el lado originalmente no dado, $AL$, subtienda un ángulo recto a cada uno de los vértices no pertenecientes a él, en $B$, $C$, ... $F$, ... $K$.* Si hay un área máxima ésta debe ser alcanzada en la situación descrita. Podemos tomar por garantía que hay un área máxima y, recordando un poco de geometría elemental, describir la situación en otros términos, como sigue: *el máximo del área es alcanzado si, y sólo si, el polígono es inscrito en un semicírculo, cuyo diámetro es el lado originalmente no dado.*

Hemos obtenido aquí el mismo resultado que en la sección 5 (3), aunque ahora no hemos usado el teorema isoperimétrico y allí sí lo hicimos.

(3) Hemos verificado primero, en (1), una consecuencia muy particular del teorema isoperimétrico; después, en (2), una consecuencia mucho más amplia. Quizá tenemos ya suficientes elementos para atacar otra consecuencia más amplia, derivada arriba en la sección 5 (2).

Comparemos dos polígonos $ABC \ldots KL$ y $A'B'C' \ldots K'L'$; ver la figura 10.10. Los lados correspondientes son iguales $AB = A'B'$, $BC = B'C' \ldots$, $KL = K'L'$, $LA = L'A'$, pero algunos de los ángulos son diferentes; $ABC \ldots KL$ está inscrito en un círculo, pero $A'B'C' \ldots K'L'$ no.

Unamos un vértice $J$ de $ABC \ldots KL$ al centro del círculo circunscrito y dibujemos el diámetro $JM$. Si, por casualidad, el punto $M$ coincide con un vértice de $ABC \ldots KL$, nuestra tarea se ve grandemente simplificada (podríamos entonces utilizar el resultado de (2) inmediatamente). Si no, $M$ estará en el círculo entre dos vértices adyacentes del polígono incrito, digamos $A$ y $B$. Unamos $MA$, $MB$, consideremos $\Delta AMB$ (sombreado en la figura 10.10) y construyamos sobre la base $A'B'$ el $\Delta A'M'B'$ (también sombreado) congruente con $\Delta AMB$. Finalmente, unamos $J'M'$.

El polígono $AMBC \ldots KL$ está dividido en dos partes por la línea $JM$ (ver figura 10.10; el polígono que le corresponde está correspondientemente dividido por $J'M'$). Aplicar a ambas partes el teorema probado en (2). El área del polígono $MBC \ldots J$, inscrito en un semicírculo, no es menor que el área de $M'B'C' \ldots J'$; en efecto, los lados correspondientes son todos iguales, excepto $MJ$, que forma el diámetro del semicírculo, el cual puede diferir de $M'J'$. Por la misma razón, el área de $MALK \ldots J$ no es menor que el de $M'A'L'K' \ldots J'$. Por adición obtenemos que

$$\text{área } AMBC \ldots KL > \text{ área } A'M'B'C' \ldots K'L'.$$

Sin embargo,

$$\Delta AMB \cong \Delta A'M'B'.$$

Por sustracción obtenemos que

$$\text{área } ABC \ldots KL > \text{área } A'B'C' \ldots K'L'.$$

*El área de un polígono inscrito en un círculo es más grande que el área de otro polígono con los mismos lados.*

Hemos obtenido aquí el mismo resultado que en la sección 5 (2), pero sin usar el teorema isoperimétrico que utilizamos allí.

(La primera desigualdad, entre las áreas de los polígonos extendidos,

contiene el signo $>$, aunque un lector consciente podría haber esperado el signo $\geqq$. Añadimos la discusión de este, en cierta manera, sutil punto. Decimos que el polígono $A'M'B'C' \ldots K'L'$ no es inscribible

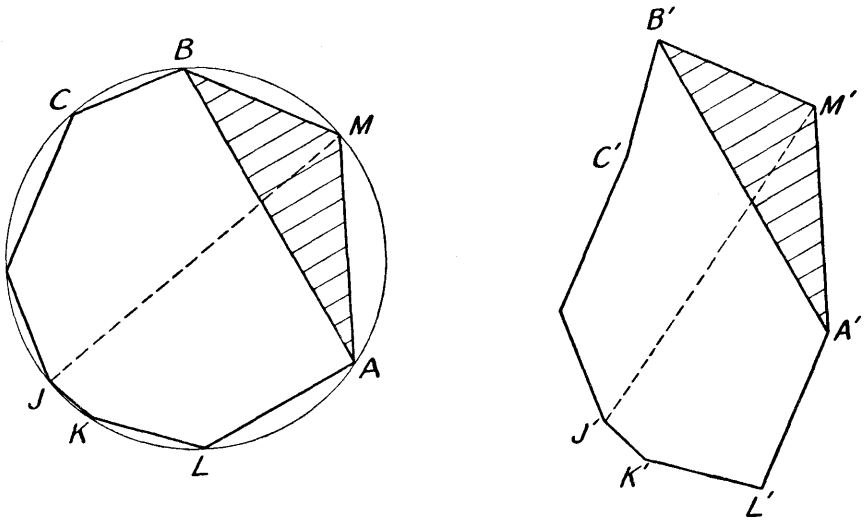



Fig. 10.10. Un polígono está inscrito; el otro, no.

en un círculo; de otro modo, $A'B'C' \ldots K'L'$ sería también inscribible, lo que no ocurre. Decimos también que los polígonos $M'B'C' \ldots J'$ y $M'A'L'K' \ldots J'$ no son ambos inscribibles en un semicírculo con diámetro $M'J'$; de otro modo, el polígono total $A'M'B'C' \ldots K'L'$ sería inscribible en un círculo, lo que no ocurre. De aquí las palabras «no menor», usadas dos veces en la derivación de la desigualdad en cuestión, pueden ser reemplazadas al menos una vez por «más grande» *).

## 7. Muy cerca

Las consecuencias que hemos logrado verificar hacen el teorema isoperimétrico altamente plausible. Aún es más. Tenemos el sentimiento de que estas consecuencias «contienen un leto», que estamos «muy cerca» de la solución final, de la prueba completa.

(1) *Encontrar el polígono que con un número dado de lados y un perímetro dado tiene el área máxima.*

Si hay un polígono tal debe estar *inscrito en un círculo*. Esto podemos concluir de nuestra última observación, sección 6 (3).

Por otro lado, tenemos el problema casi resuelto. Asumimos que conocemos ya la posición de todos los vértices menos uno, por ejemplo,

* Los teoremas y demostraciones de esta sección se deben a Lhuilier; ver nota a pie de página del capítulo VIII.

$X$. Luego otros $n-1$ vértices, digamos, $U, \ldots, W, Y$ y $Z$, están ya fijos. El polígono total $U \ldots WXYZ$ consta de dos partes: el polígono $U \ldots WYZ$ con $n-1$ vértices fijos, que es independiente de $X$, y $\Delta WXY$, que depende de $X$. De este triángulo, $\Delta WXY$, conocemos la base $WY$ y la suma de los otros dos lados $WX + XY$; en efecto, los lados $n-2$ que quedan del polígono suponemos que son conocidos y que, en realidad, conocemos la suma de todos los $n$ lados. El área de $\Delta XWY$ debe ser un máximo. Sin embargo, como resulta inmediato, el $\Delta WXY$ con base y perímetro conocido alcanza su área máxima cuando es isósceles (ejemplo 8.8). Es decir, $WX = XY$, los dos lados adyacentes del polígono buscado son iguales. Por tanto (según la simetría de las condiciones y el patrón de variación parcial), cualesquiera dos lados adyacentes son iguales. Todos los lados son iguales: el polígono buscado es *equilátero.*

El polígono buscado, que es equilátero y está inscrito en un círculo, es necesariamente regular: *De todos los polígonos con un número dado de lados y un perímetro también dado el polígono regular es el que tiene área más grande.*

(2) *Dos polígonos regulares, uno con n lados y otro con* $n+1$ *lados, tienen el mismo perímetro. ¿Cuál de ellos es el de área mayor?*

El polígono regular de $n+1$ lados tiene un área mayor que todos los polígonos irregulares con $n+1$ lados y el mismo perímetro, como hemos visto en (1). Sin embargo, el polígono regular con $n$ lados, cada uno igual a $a$, por ejemplo, puede ser considerado como un polígono irregular con $n+1$ lados: $n-1$ lados con de longitud $a$, dos de longitud $a/2$ y un ángulo es igual a 180°. (Consideremos el punto medio de un lado del polígono, concebido a la manera usual, como un vértice, y después llegamos a la presente concepción menos usual.) Por tanto, *el polígono regular con* $n+1$ *más lados tiene un área más larga que el polígono regular con n lados y el mismo perímetro.*

(3) *Un círculo y un polígono regular tienen el mismo perímetro. ¿Cuál es el de área mayor?*

Démonos cuenta de lo que el resultado anterior, en (2), significa. Tomemos $n = 3, 4, \ldots$ y establezcamos el resultado en cada caso particular. Al pasar de un triángulo equilátero a un cuadrado con el mismo perímetro encontramos el área aumentada. Al pasar de un cuadrado a un pentágono regular con el mismo perímetro, de nuevo encontramos el área aumentada. Y así sucesivamente, pasando de una figura regular a la próxima, de un pentágono a un hexágono, de un hexágono a un heptágono, de $n$ a $n+1$, vemos que el área aumenta a cada paso cuando el perímetro permanece invariable. Al fin, en el límite, encontra-

mos el círculo. Su perímetro es aún el mismo, pero su área es obviamente superior al área de cualquier polígono regular en la secuencia infinita de la cual él es el límite. *El área del círculo es mayor que la de cualquier polígono regular con el mismo perímetro.*

(4) *Un círculo y un polígono arbitrario tienen el mismo perímetro. ¿Cuál tiene el área más grande?*

El círculo, como se sigue inmediatamente de (1) y de (2).

(5) *Un círculo y una curva arbitraria tienen el mismo perímetro. ¿Cuál tiene el área mayor?*

El círculo, como se sigue de (4), puesto que cualquier curva es el límite de polígonos. ¡Hemos probado el teorema isoperimétrico!

## 8. Tres formas del teorema isoperimétrico

En lo que antecede (secciones 6 y 7) hemos probado el siguiente enunciado del teorema isoperimétrico:

I. *De todas las figuras planas de igual perímetro el círculo es la de área máxima.*

En la sección 2, no obstante, hemos discutido otro enunciado.

II. *De todas las figuras planas de igual área el círculo es la de perímetro mínimo.*

Estos dos enunciados son diferentes, y no sólo difieren en palabras. Necesitan mayor clarificación.

(1) Dos curvas se llaman «isoperimétricas» si sus perímetros son iguales. «De todas las curvas planas isoperimétricas el círculo es la de área mayor»; éstas son las palabras tradicionales del enunciado I, que explica el nombre de «teorema isoperimétrico».

(2) Podemos llamar a los dos enunciados del teorema (I y II) «enunciados conjugados» (ver sección 8.6). Mostraremos que estos dos enunciados conjugados son equivalentes el uno al otro, mostrando que ambos son equivalentes a un tercero.

(3) $A$ denota el área y $L$ la longitud del perímetro de una curva dada. Asumimos que la curva dada y un círculo de radio $r$ son isoperimétricos: $L = 2\pi r$. Luego la primera forma (enunciado I) del teorema isoperimétrico sostiene que

$$A \leqq \pi r^2.$$

Sustituyendo $r$ en términos de $L$, $r = L/2\pi$, podemos fácilmente transformar la desigualdad en esta otra

$$\frac{4\pi A}{L^2} \leqq 1.$$

Llamamos a esta desigualdad la *desigualdad isoperimétrica* y al cociente del lado izquierdo *cociente isoperimétrico*. Este cociente depende sólo de la forma de la curva y es independiente de su tamaño. En efecto, si, sin cambiar la forma, alargamos las dimensiones lineales de la curva de radio 1 : 2, el perímetro alcanza $2L$ y el área $4A$, mientras el cociente $A/L^2$ permanece igual, y lo mismo es cierto de $4\pi A/L^2$ y del alargamiento de cualquier radio. Algunos autores llaman a $A/L^2$ el cociente isoperimétrico; hemos introducido el factor $4\pi$ para hacer nuestro cociente isoperimétrico igual a 1 en el caso del círculo. Con esta terminología podemos decir:

III. *De todas las figuras planas el círculo tiene el cociente isoperimétrico más alto* *.

Esta es la tercera forma del teorema isoperimétrico.

(4) Hemos llegado a la tercera forma del teorema viniendo de figuras con igual perímetro. Ahora empecemos con el enunciado III y pasemos a las figuras con igual área. Asumimos que una curva con área $A$ y perímetro $L$ tiene el mismo área que un círculo con radio $r$. Es decir, $A = \pi r^2$. Sustituyendo $A$ podemos transformar la desigualdad isoperimétrica en $L \geq 2\pi r$. Es decir, el perímetro de la figura es más grande que el del círculo de área igual. Hemos llegado así a la segunda forma del teorema, en el enunciado II.

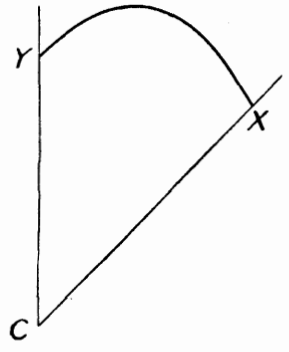



Fig. 10.11. El problema de Dido complicado con un cabo.

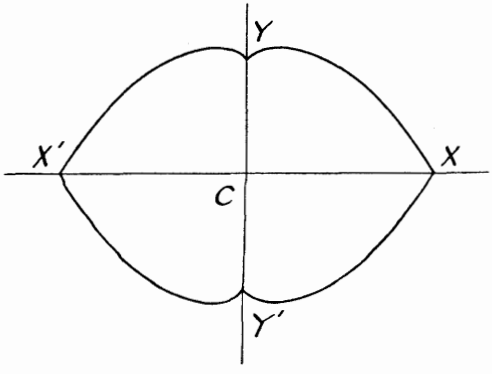



Fig. 10.12. La reflexión lo resuelve algunas veces.

(5) Podríamos, por supuesto, proceder en la misma línea argumental, aunque, en dirección opuesta y pasando por III, derivar I de II. Y así podemos estar satisfechos de que las tres formas sean equivalentes.

* Si abreviamos «cociente isoperimétrico» por C. I. podríamos decir que el círculo tiene el C. I. más alto.

## 9. Aplicaciones y problemas

Si Dido pactó con los nativos en la cercanía de un cabo, su problema fue, quizá, más semejante al que sigue que al discutido en la sección 5 (1).

*Dado un ángulo* (la parte infinita de un plano entre dos rayos trazados desde el mismo punto inicial). *Encontrar el área máxima sacada de éste por una línea de longitud dada.*

En la figura 10.11 el vértice del ángu'o dado es llamado $C$ (cabo). La línea arbitraria que conecta los puntos $X$ e $Y$ se supone que tiene la longitud dada $l$. Se nos pide hacer un máximo del área de tres esquinas entre la curva y el litoral. Podemos deslizar los puntos finales $X$ e $Y$ de la curva y modificar su forma, pero no podemos cambiar su longitud $l$.

El problema no es demasiado fácil, pero es uno de esos problemas que una elección particular de los datos lo hace más accesible. Si el ángulo en $C$ es un ángulo recto podemos tomar la imagen del espejo primero con respecto a un lado del ángulo, luego con respecto al otro. Obtenemos así una nueva figura, la 10.12, y un nuevo problema. La línea $XY$, cuadruplicada por reflexiones, nos proporciona una nueva línea *cerrada* de longitud dada, $4l$. El área, para ser maximizada, cuadruplicada por reflexiones, nos proporciona una nueva área que ha de ser maximizada, *completamente rodeada* por la nueva curva dada. Según el teorema isoperimétrico, la solución del nuevo problema es un círculo. Este círculo tiene dos ejes de simetría dados, $XX'$ e $YY'$, y así su centro está en la intersección de estos dos ejes, en el punto $C$. Por tanto, la solución del problema original (el problema de Dido) es un *cuadrante:* un cuarto de un círculo con centro en el vértice del ángulo dado.

Aquí, naturalmente, recordamos la solución de la sección 5 (1) basado en la figura 10.2 y observamos que es estrechamente análoga a la solución presente. Es fácil ver que hay una infinidad de casos en que soluciones de esta clase valen. Si el ángulo dado en $C$ es $360°/2n = 180°/n$, podemos transformar, por reflexiones repetidas, la curva $XY$ con longitud $l$ dada en una nueva y próxima curva de longitud $2nl$ y el problema propuesto en un nuevo problema cuya solución es un círculo, por virtud del teorema isoperimétrico. El caso tratado en la sección 5 (1) y en la presente sección son sólo los dos primeros casos en esta secuencia infinita, que corresponden a $n = 1$ y $n = 2$.

Esto es, si el ángulo en $C$ es de una clase especial ($180°/n$ con integral $n$), la solución de nuestro problema (figura 10.11) es un arco circular con centro en $C$. Es natural esperar que esta forma de solución es independiente de la magnitud del ángulo (al menos mientras no exceda

los 180°). Es decir, conjeturamos que la solución del problema de la figura 10.11 es el arco de un círculo con centro en $C$, tanto si el ángulo en $C$ es o no de la clase especial $180°/n$. Esta conjetura es inductiva, apoyada por la evidencia de una infinidad de casos particulares, $n = 1, 2, 3, \ldots$ ¿Es esta conjetura cierta?

La anterior aplicación del teorema isoperimétrico y los problemas relacionados pueden hacernos anticipar muchas aplicaciones y problemas semejantes. Nuestra derivación del teorema levanta más cuestiones; sus análogos en geometría sólida y física matemática sugieren todavía otras cuestiones más. El teorema isoperimétrico, profundamente arraigado en nuestra intuición y experiencia, tan fácil de conjeturar, aunque no de probar, es una inagotable fuente de inspiración.

## EJEMPLOS Y COMENTARIOS

### Primera parte

**1.** *Mirando atrás.* En lo anterior (secciones 6-8) hemos probado el teorema isoperimétrico —¿lo hemos?—. Comprobemos el argumento paso a paso.

En ello no parece haber objeción contra el sencillo resultado de la sección 6 (1). Sin embargo, al resolver el problema de la sección 6 (2) asumimos la existencia del máximo sin prueba; y volvimos a hacer lo mismo en la sección 7 (1). ¿Invalidan el resultado estas improbadas asunciones?

**2.** *¿Podríamos derivar parte del resultado de modo diferente?* Verificar el caso particular más sencillo y no-trivial del resultado encontrado directamente en la sección 5 (2). Es decir, probar independientemente de la sección 6 (3) que el área de un cuadrilátero inscrito en un círculo es más grande que el área de cualquier otro cuadrilátero con los mismos lados. [Ejemplo 8.41.]

**3.** *Volver a formular con más detalle* el argumento de la sección 7 (2): construir un polígono con $n + 1$ lados que tiene el mismo perímetro, aunque mayor área, que el polígono regular con $n$ lados.

**4.** Probar independientemente de la sección 7 (3) que un círculo tiene un área más grande que cualquier polígono regular con el mismo perímetro.

**5.** Probar, más generalmente, que un círculo tiene un área mayor que cualquier otro polígono circunscribible con el mismo perímetro.

**6.** Volver a formular con más detalle el argumento de la sección 7 (5). ¿Prueba éste el enunciado I de la sección 8? ¿Hay alguna objeción?

**7.** *¿Podemos utilizar el método para algún otro problema?* Utilizar el método de la sección 8 para probar que los dos siguientes enunciados son equivalentes:

«De todas las cajas con una superficie dada, el cubo es la de máximo volumen».

«De todas las cajas con un volumen dado, el cubo tiene la superficie mínima».

**8.** *Formas perspicaces del teorema isoperimétrico.* Comparar los enunciados I, II y III de la sección 8 con los siguientes:

I′. *El área de un círculo es más grande que cualquiera otra curva plana con el mismo perímetro.*

II′. *El perímetro de un círculo es más corto que el de cualquiera otra curva plana con el mismo área.*

III′. *Si $A$ es el área de una curva plana y $L$ la longitud de su perímetro, entonces*

$$\frac{4\pi A}{L^2} \leq 1$$

*y la igualdad es alcanzada si, y sólo si, la curva es un círculo.*

Mostrar que I′, II′ y III′ son equivalentes una a otra. ¿Hemos probado I′?

**9.** Dada una curva $C$ con perímetro $L$ y área $A$; $C$ no es un círculo. Construir una curva $C'$ con el mismo perímetro $L$, pero con un área $A'$ más grande que $A$.

Este problema es importante (¿por qué?), aunque no demasiado fácil. Si no puede resolverlo en general, resuélvalo en significativos casos particulares; propóngase problemas pertinentes que puedan acercarnos a su solución general; intentemos volver a formularlo; intentemos aproximarnos por un camino o por otro.

**10.** Es dado un cuadrilátero $C$ con ángulo reentrante, el perímetro $L$ y el área $A$. Construir un triángulo $C'$ con el mismo perímetro $L$, pero con un área $A'$ mayor que $A$.

**11.** Generalizar el ejemplo 10.

**12.** La información «$C$ no es un círculo» es «puramente negativa». ¿Podríamos caracterizar $C$ más «positivamente» de manera que nos diera un asidero para acometer el ejemplo 9?

[Tres puntos cualesquiera en una curva están en el mismo círculo, o en una línea recta. ¿Qué pasará con cuatro puntos?]

**13.** Es dada una curva $C$ con perímetro $L$ y área $A$; hay cuatro puntos $P, Q, R$ y $S$ de $C$ que no están en el mismo círculo ni en la misma línea recta. Construir una curva $C'$ con el mismo perímetro $L$, pero con un área $A'$ mayor que $A$. [Ejemplo 2.]

**14.** Comparar los dos siguientes problemas.

Consideremos curvas con un perímetro dado. Si $C$ es una curva tal, aunque no un círculo, podemos construir otra curva $C'$ con un área más grande. (En efecto, esto ha sido hecho en los ejemplos 10-13. La condición de que $C$ no es un círculo es esencial; nuestra construcción fracasa en aumentar el área del círculo.) ¿Podemos concluir de aquí que el círculo tiene el área más grande?

Consideremos enteros positivos. Si $n$ es un entero tal, aunque no 1, podemos construir otro entero $n'$ más grande que $n$. (En efecto, pongamos $n' = n^2$. La condición $n > 1$ es esencial; nuestra construcción fracasa para $n = 1$ y para $1^2 = 1$.) ¿Podemos concluir de aquí que 1 es el entero más grande?

Expresemos la diferencia, si es que hay alguna.

**15.** Probar el enunciado I′ del ejemplo 8.

## Segunda parte

**16.** *La cuerda y el palo.* Dados una cuerda y un palo, cada extremo de uno atado al correspondiente extremo del otro (por supuesto, que la cuerda debe ser más larga que el palo). Acotemos con este fin el área más larga posible.

Coloquemos el palo abajo. Sus puntos finales $A$ y $B$ determinan su posición completamente. Pero la cuerda puede cambiar infinitamente de forma, formando una curva arbitraria de longitud dada que empieza en $A$ y termina en $B$; ver la figura 10.13. Una de las posibles figuras de la cuerda es un arco circular que forma con el palo un segmento de un círculo. Completar el círculo por adición de otro segmento (sombreado en la figura 10.14, I) y añadir el mismo segmento a la figura formada por el palo y una posición arbitraria de la cuerda (figura 10.14, II). El círculo I de la figura 10.14 tiene un área mayor que

cualquiera otra curva con el mismo perímetro, y el II de la misma figura *es* una curva tal. Sustrayendo el mismo segmento (sombreado) de I y II encontramos el resultado: el área acotada por el palo y la cuerda es un *máximo cuando la cuerda forma un arco circular.* Este resultado permanece válido si añadimos al área variable de la figura 10.13 cualquier área invariable a lo largo de la línea recta invariable que la limita. Esta observación es útil con frecuencia.

Formular el resultado conjugado. Es decir, formular el hecho que tiene la misma relación al teorema II de la sección 8 como el hecho encontrado en lo anterior con respecto al teorema I de la sección 8.

**17.** Dado un ángulo (la parte infinita de un plano entre dos rayos trazados desde el mismo punto inicial) y dos puntos, uno a cada lado

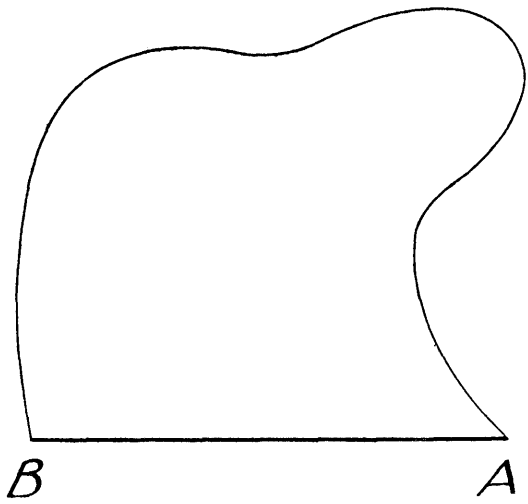



Fig. 10.13. Palo y cuerda.

del ángulo. Encontrar el área máxima cortada a partir del ángulo por una línea de longitud dada que conecta los dos puntos dados. (En la figura 10.11 los puntos $X$ e $Y$ estan dados).

**18.** Dado un ángulo con una apertura menor de 180° y un punto sobre uno de sus lados, encontrar el área máxima cortada a partir del ángulo por una línea de longitud dada que empieza en el mismo punto. (En la figura 10.11 el punto $X$ es dado, pero $Y$ es variable.)

**19.** Dado un ángulo con una apertura menor de 190°, encontrar el área máxima cortada a partir del ángulo por una línea de longitud

dada. (En la figura 10.11 los puntos $X$ e $Y$ son variables. Una conjetura fue establecida en la sección 9.)

**20.** Dado un ángulo con una apertura menor de 180°, encontrar el área máxima cortada a partir del ángulo por una línea recta de longitud dada.

**21.** *Dos palos y dos cuerdas.* Tenemos dos palos, $AB$ y $CD$. Una primera cuerda está atada al último punto $B$ del primer palo con un cabo y al primer punto $C$ del segundo palo con el otro cabo. Otra cuerda conecta de modo semejante $D$ y $A$. Acotar de esta manera el mayor área posible.

**22.** Generalizar.

**23.** Especializar, y obtener por especialización un teorema elemental que ha jugado un papel importante en el texto.

**24.** Es dado un círculo en el espacio. Encontrar la superficie con área dada limitada por el círculo dado que forma el volumen máximo con el disco ceñido por el círculo dado. [¿Conocemos un problema análogo?]

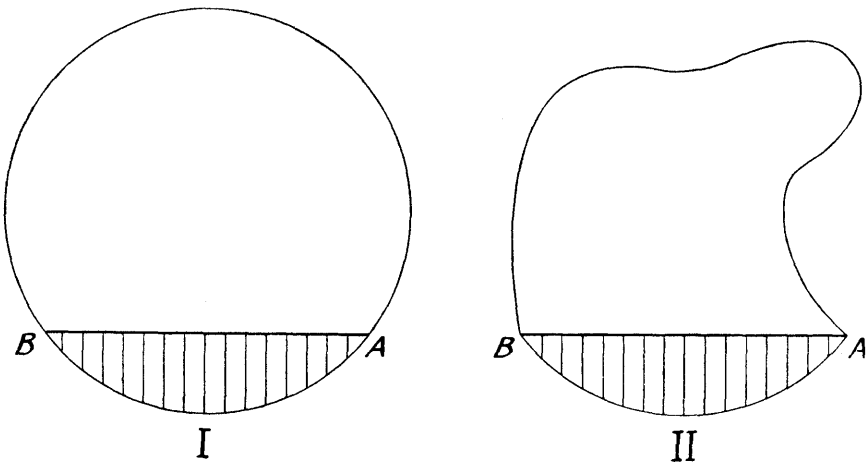



Fig. 10.14. El principio del arco circular.

**25.** *El problema de Dido en geometría sólida.* Dado un ángulo triedro (una de las ocho partes infinitas en que el espacio es dividido por tres planos que intersectan en un punto). Encontrar el volumen máximo cortado a partir del ángulo triedro por una superficie de área dada.

Este problema es demasiado difícil. Se pide solamente extraer un caso particular más accesible.

**26.** Encontrar un problema análogo al ejemplo 25, cuyo resultado podemos predecir. [Generalizar, especificar, pasar el límite, ...]

**27.** *Bisectores de una región del plano.* Consideremos la región de un plano ceñida por una curva. Un arco que una dos puntos de la curva que lo rodea se llama *bisector* de la región siempre que divida la región en dos partes de área igual.

Mostrar que dos bisectores cualesquiera de la misma región tienen por lo menos un punto común.

**28.** Comparar dos bisectores de un cuadrado. Uno de ellos es una línea recta paralela a uno de los lados que pasa por el centro del cuadrado. El otro es un cuarto de un círculo cuyo centro es un vértice. ¿Cuál de los dos es más corto?

**29.** Encontrar el bisector recto más corto de un triángulo equilátero.

**30.** Encontrar el bisector más corto de un triángulo aquilátero.

**31.** Mostrar que los bisectores más cortos de un círculo son sus diámetros.

**32.** Encontrar el bisector más corto de una elipse.

**33.** Intentar formular teoremas generales que cubran los ejemplos 28-32.

**34.** *Bisectores de una superficie cerrada* *. Una curva cerrada que no se autointersecta se llama *bisector* de la superficie siempre que divida la superficie en dos partes (superficies abiertas) de área igual.

Mostrar que dos bisectores cualesquiera de la misma superficie tienen al menos un punto común.

**35.** Un bisector más corto de la superficie de un poliedro consta de piezas cada una de las cuales es o una línea recta o un arco circular.

**36.** Un bisector más corto de la superficie de un sólido regular es un polígono regular. Encontrar su forma y situación, y el número de soluciones para cada uno de los cinco sólidos regulares. (Podemos experimentar con un modelo del sólido y una banda de goma.)

**37.** Mostrar que los bisectores más cortos de una superficie esférica son los círculos grandes.

* Consideramos aquí sólo superficies cerradas del «tipo topológico» de la esfera y excluir, por ejemplo, el toro (de la forma de buñuelo).

**38.** Intentar encontrar una generalización del ejemplo 37 que cubra también una parte sustancial del ejemplo 36. [Ejemplos 9.23, 9.24.]

**39.** Es dada una esfera $S$ con radio $a$. Llamamos *diafragma* de $S$ la parte de una superficie esférica que intersecta $S$ que está dentro de $S$. Probar:

(1) Todos los diafragmas que pasan por el centro de $S$ tienen el mismo área.

(2) Ningún diafragma que biseccione el volumen de $S$ tiene un área menor que $\pi a^2$.

El último enunciado, y los casos análogos discutidos, sugieren una conjetura. Establecerla. [Ejemplos 31, 37.]

**40.** *Una figura de muchas perfecciones.* Consideremos una región de un plano rodeada por una curva. Deseamos examinar alguno de los muchos teoremas análogos al teorema isoperimétrico: De todas las regiones con área dada el círculo tiene el perímetro mínimo.

Antes nos encontramos con un teorema de esta clase. En la sección 4 consideramos alguna evidencia inductiva para el enunciado: De todas las membranas con un área dada la membrana circular emite el tono principal más profundo.

Ahora consideremos la región como un plato homogéneo de espesor uniforme. Miraremos el momento de inercia de este plato sobre un eje perpendicular a él y que pasa por su centro de gravedad. Este momento de inercia, que llamamos el «momento polar de inercia», depende, siendo igual en otras cosas, del tamaño y la forma del plato. De todos los platos con un área dada el plato circular tiene el momento polar de inercia mínimo.

Este plato, si es conductor de electricidad, puede también recibir una carga eléctrica, proporcional a su capacidad electrostática. También la capacidad depende de la forma y tamaño del plato. De todos los platos con un área dada el plato circular tiene la capacidad mínima.

Ahora sea la región una sección transversal de una viga homogénea y elástica. Si retorcemos la viga sobre su eje, podemos observar que lo resiste muy bien. Esta resistencia, o «rigidez torsional», de la viga depende, siempre que otras cosas sean iguales, del tamaño y la forma de la sección transversal. De todas las secciones transversales con área dada la circular tiene la máxima rigidez torsional *.

---

* Para las pruebas de los teoremas indicados y de otros semejantes ver G. POLYA y G. SZEGÖ, *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, 1951.

¿Por qué es el círculo la solución de tantos y tan diferentes problemas de máximos y mínimos? ¿Cuál es la «razón»? ¿Es la «perfecta simetría» del círculo la «verdadera razón»? Tales vagos problemas pueden ser fecundos y estimulantes, siempre que no caigamos en mera charla y especulación superflua, sino que intentemos seriamente alcanzar algo más preciso y concreto.

**41.** *Un caso análogo.* ¿Vemos la analogía entre el teorema isoperimétrico y el teorema de los medios? (Véase sección 8.6.)

La longitud de una curva cerrada depende de la misma manera de cada punto, o de cada elemento, de la curva. También el área de la región rodeada por la curva depende de la misma manera de cada punto, o elemento, de la curva. Buscamos el máximo del área cuando la longitud es dada. Como las dos cantidades afectadas son de tal naturaleza que ningún punto de la curva juega un papel favorable en su definición, no debe sorprendernos que la solución sea la única curva cerrada que contiene cada uno de sus puntos en la misma manera, y cualesquiera dos elementos de la misma son superponibles: el círculo.

La suma $x_1 + x_2 + \ldots + x_n$ es una función simétrica de las variables $x_1, x_2, \ldots x_n$; es decir, que depende del mismo modo en cada variable. También el producto $x_1, x_2 \ldots x_n$ depende igualmente de cada variable. Buscamos el máximo del producto cuando la suma es dada. Como las dos cantidades afectadas son simétricas en las $n$ variables, no debemos sorprendernos de la solución $x_1 = x_2 = \ldots = x_n$.

Además del área y la longitud hay otras cantidades que dependen del tamaño y la forma de una curva cerrada, que «depende igualmente de cada elemento de la curva»; hemos visto varias cantidades tales en el ejemplo 40. Busquemos el máximo de una cantidad de esta clase cuando otra cantidad de la misma clase es dada. ¿Es la solución, si es que hay alguna, necesariamente el círculo?

Retornemos, para una contestación plausible, al más sencillo caso análogo. Consideremos dos funciones simétricas, $f(x_1, x_2, \ldots x_n)$ y $g(x_1, x_2 \ldots x_n)$, de $n$ variables y busquemos los extremos de $f(x_1, x_2 \ldots x_n)$ cuando nos es dado que $g(x_1, x_2, \ldots x_n) = 1$. Se dan casos en que no hay máximo, otros en que no hay mínimo y otros aún en que no hay máximos ni mínimos. La condición $x_1 = x_2 = \ldots = x_n$ juega un papel importante *, aunque no estemos satisfechos cuando un

* G. H. Hardy, J. E. Littlewood y G. Polya. *Inequalities*. Véase págs. 109-110 y los teoremas allí citados.

máximo o un mínimo es alcanzado. Hay, no obstante, un hecho sencillo: si

$$x_1 = a_1, \qquad x_2 = a_2, \qquad x_3 = a_3 \ \ldots \ x_n = a_n$$

es una solución, también

$$x_1 = a_2, \qquad x_2 = a_1, \qquad x_3 = a_3 \ \ldots \ x_n = a_n$$

es una solución, por la simetría de las funciones $f$ y $g$. Por tanto, si $a_1 \neq a_2$, hay por lo menos dos soluciones diferentes. *Si hay una solución única* (es decir, si el extremo es alcanzado, y alcanzado por un sistema de valores de $x_1, x_2 \ldots x_n$) *la solución requiere* $x_1 = x_2 = \ldots = x_n$.

«Comparaison n'est pas raison», dicen los franceses. Por supuesto, que comparaciones como la precedente no pueden darnos una razón válida, sino sólo una indicación heurística. Sin embargo, podemos agradecer muchas veces el recibir una indicación semejante.

Tomemos una ilustración

$$f(x_1, x_2, \ldots x_n) = (x_1 + x_2 + \ldots + x_n)^2,$$
$$g(x_1, x_2, \ldots x_n) = (x_1^2 + x_2^2 + \ldots x_n^2)/n,$$

y encontremos los extremos de $f$ bajo la condición de que $g = 1$, considerando: (1) todos los valores reales de $x_1, x_2, \ldots x_n$; y, (2) sólo los valores reales no-negativos de estas variables.

**42.** *Los sólidos regulares.* Encontrar el poliedro con un número dado $n$ de caras y con un área-superficie dada que tenga el volumen máximo.

Este problema muy difícil es sugerido por los problemas análogos de la sección 7 (1), que a su vez sugiere una conjetura: si hay un sólido regular con $n$ caras, éste nos da el volumen máximo. Por muy plausible que esta conjetura pueda parecer, no obstante la encontramos equivocada en dos casos de cinco. En efecto, la conjetura es

correcta para $n = 4, 6, 12,$

incorrecta para $n = 8, 20.$

¿Cuál es la diferencia? Intentemos observar alguna sencilla propiedad geométrica que distingue entre las dos clases de sólidos regulares.

**43.** *Razones inductivas.* $V$ denota el volumen de un sólido y $S$ el área de su superficie. Por analogía, la sección 8 (3) sugiere definir

$$\frac{36\pi V^3}{S^2}$$

como el cociente isoperimétrico en geometría sólida. Por analogía, podemos conjeturar que la esfera tiene el cociente isoperimétrico más alto. La tabla III apoya inductivamente esta conjetura.

Comprobar algunas de las figuras dadas en la tabla III y añadir nuevo material. En especial, tratar de encontrar un sólido con un cociente isoperimétrico más alto que el del icosaedro regular.

TABLA III

EL COCIENTE ISOPERIMETRICO $36\pi V^2/S^3$

| | |
|---|---|
| Esfera | 1,0000 |
| Icosaedro | 0,8288 |
| El mejor cono doble | 0,7698 |
| Dodecaedro | 9,7547 |
| El mejor prisma | 0,6667 |
| Octaedro | 0,6045 |
| Cubo | 0,5236 |
| El mejor cono | 0,5000 |
| Tetraedro | 0,3023 |

Para el «mejor» cono doble, prisma y cono véanse los ejemplos 8.38, 8.35 y 8.52, respectivamente.

# Más clases de razones plausibles

*Las relaciones más sencillas son las más comunes, y éste es el fundamento en que descansa la inducción.*—LAPLACE *.

## 1. CONJETURAS Y CONJETURAS

Todas nuestras discusiones anteriores tratan sobre el papel de las conjeturas en la investigación matemática. Nuestros ejemplos nos dieron oportunidad de familiarizarnos con dos clases de argumentos plausibles en favor o en contra de una consecuencia propuesta: discutimos argumentos inductivos, a partir de la verificación de consecuencias o a partir de la analogía. ¿Hay otra clase de argumentos plausibles útiles en favor o en contra de una conjetura? Los ejemplos de este capítulo tratan de aclararnos esta cuestión.

También nos daremos cuenta de que hay conjeturas de varias clases: grandes y pequeñas, originales y rutinarias. Hay conjeturas que juegan un papel espectacular en la historia de la ciencia, pero también las soluciones de los problemas matemáticos más modestos necesitan una conjetura o intuición igualmente modesta. Empecemos con ejemplos discutidos en el aula y después continuemos con otros de importancia histórica.

## 2. JUZGANDO POR UN CASO AFÍN

Al trabajar en un problema, a menudo intentamos intuir. Por supuesto que nos gustaría intuir la solución total. Si, no obstante, no tenemos éxito en esto, nos sentimos bastante satisfechos si podemos intuir uno u otro aspecto de la solución. Al menos, nos gustaría saber si nuestro problema es «razonable». Preguntémonos: ¿Es razonable nuestro problema? *¿Es posible satisfacer la condición? ¿Es suficiente la condición para determinar la incógnita? ¿O es insuficiente? ¿O redundante? ¿O contradictoria?* **.

---

* *Essai philosophique sur les probabilités;* ver *Oeuvres Complètes de Laplace,* vol. 7, página CXXXIX.

** *How to Solve It,* pág. 111.

Tales cuestiones se nos presentan naturalmente, y son muy útiles en un primer estadio de nuestro trabajo, cuando aún no necesitan una contestación definitiva, sino sólo provisional, es decir, cuando la contestación es una intuición; y hay casos en que podemos intuir la solución bastante razonablemente y con muy poca duda.

Para ilustrar esto consideremos un problema elemental en geometría sólida. El eje de un cilindro pasa por el centro de una esfera. La superficie del cilindro intersecta la superficie de la esfera y divide la esfera sólida en dos porciones: la «esfera perforada» y la «cuña». La primera porción es exterior al cilindro y la segunda interior. Ver la figura 11.1, que gira sobre la línea vertical $AB$. *Dado* $r$, *el radio de la esfera, y* $h$, *la altura del agujero cilíndrico, encontrar el volumen de la esfera perforada.*

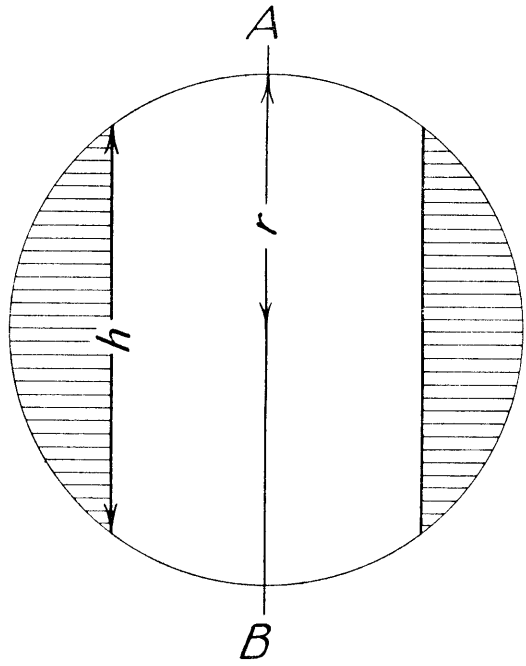



Fig. 11.1. La esfera perforada.

Al familiarizarnos con el problema propuesto llegamos naturalmente a las cuestiones usuales: *¿Son estos datos suficientes para determinar la incógnita? ¿O son insuficientes? ¿O redundantes?* Los datos $r$ y $h$ parecen ser bastantes. En efecto, $r$ determina el tamaño de la esfera y $h$ el del agujero cilíndrico. Conociendo $r$ y $h$ podemos determinar el tamaño y la forma de la esfera perforada.

Sin embargo, al calcular el volumen requerido encontramos que

es igual a $\pi h^3/6$; véase ejemplo 5. Este resultado parece extremadamente paradójico. Nos hemos convencido de que necesitamos $r$ y $h$ para determinar la forma y el tamaño de la esfera perforada y ahora encontramos que no necesitamos $r$ para determinar su volumen; esto parece bastante raro.

Sin embargo, no hay contradicción. Si $h$ permanece constante y $r$ aumenta, la esfera perforada cambia de forma considerablemente: se hace más ancha (lo que tiende a aumentar el volumen), pero su superficie exterior se hace más llana (lo que tiende a disminuir el volumen). Sólo que nosotros no predijimos (y ello parece bastante inverosímil *a priori*) que estas dos tendencias se equilibran mutuamente y el volumen permanece igual.

Al objeto de comprender el presente caso particular y la idea general subyacente necesitamos una distinción. Debemos distinguir entre dos problemas claramente relacionados, aunque diferentes. Dados $r$ y $h$, nos pueden pedir que determinemos:

(*a*) el volumen, y

(*b*) el tamaño y la forma

de la esfera perforada. Nuestro problema original era (*a*). Hemos visto intuitivamente que los datos $r$ y $h$ son ambos necesarios y suficientes para resolver (*b*). De aquí se sigue que estos datos son también suficientes para resolver (*a*), pero no que sean necesarios para resolver (*a*); en efecto, no lo son.

Al responder a la cuestión «¿Son necesarios los datos?» *juzgamos por un caso afín, sustituimos* (*b*) por (*a*), *desdeñamos la distinción entre el problema original* (*a*) *y el problema modificado* (*b*). Desde el punto de vista heurístico se puede defender tal negligencia. Hemos necesitado sólo una respuesta provisional, pero rápida. Por otra parte, tal diferencia es *usualmente* desdeñable: los datos necesarios para determinar el tamaño y la forma son también *usualmente* necesarios para determinar el volumen. Hemos llegado a vernos envueltos en una paradoja por olvidar que nuestra conclusión era solamente heurística, o por creer de forma algo confusa que lo que no es usual nunca puede ocurrir. Y, en nuestro ejemplo, ocurrió.

*Juzgar un problema propuesto por un problema modificado* es un procedimiento heurístico razonable y que puede defenderse. No olvidemos, no obstante, que la conclusión a que llegamos por tal procedimiento es solamente provisional y no definitiva; es sólo plausible, pero de ninguna manera verdadera.

### 3. Juzgando por el caso general

Puede ser conveniente discutir el siguiente problema en una clase de álgebra para principiantes.

El testamento de un padre de tres hijos contiene las siguientes disposiciones: «La parte de mi hijo mayor será la media de las partes de los otros dos más 3.000 dólares. La parte de mi segundo hijo será exactamente la media de las partes de los otros dos. La parte del más joven será la misma menos 3.000 dólares». *¿Cuáles son las tres partes?*

*¿Es la condición suficiente para determinar las incógnitas*? Hay una buena razón para decir sí. En efecto, hay tres incógnitas; pongamos $x$, $y$ y $z$, las partes del hijo mayor, del segundo y del más joven, respectivamente. Cada una de las tres frases citadas del testamento puede ser convertida en una ecuación. Ahora bien, *en general, un sistema de tres ecuaciones con tres incógnitas determina las incógnitas.* Así, podemos pensar razonablemente del problema propuesto que es suficiente para determinar las incógnitas.

No obstante, al escribir las ecuaciones obtenemos el siguiente sistema:

$$x = \frac{y + z}{2} + 3.000$$

$$y = \frac{x + z}{2}$$

$$z = \frac{x + y}{2} - 3.000.$$

Sumando estas ecuaciones obtenemos

$$x + y + z = x + y + z,$$

o

$$0 = 0.$$

Por tanto, cualquier ecuación del sistema es una consecuencia de las otras dos ecuaciones. Nuestro sistema contiene sólo dos ecuaciones *independientes* y, por tanto, en efecto, *no* es suficiente determinar las incógnitas.

El problema se modifica esencialmente si el testamento contiene

también las siguientes frases: «Divido toda mi fortuna de 15.000 dólares entre mis tres hijos». Este sistema añade al anterior la ecuación

$$x + y + z = 15.000.$$

Tenemos ahora un sistema más comprensivo de cuatro ecuaciones. Sin embargo, *en general, un sistema de cuatro ecuaciones con tres incógnitas es contradictorio.* De hecho, no obstante, el presente sistema no es contradictorio, sino suficiente para determinar las incógnitas, y da por resultado

$$x = 7.000, \quad y = 5.000 \text{ y } z = 3.000.$$

Las contradicciones aparentes de este ejemplo no demasiado profundo tampoco son muy difíciles de desenmarañar, pero una explicación cuidadosa puede ser útil.

No es verdad que «un sistema de $n$ ecuaciones con $n$ incógnitas determine las incógnitas». En efecto, acabamos de ver un ejemplo en contra de esto con $n = 3$. Lo que importa aquí, sin embargo, no es un teorema matemático, sino una enunciación heurística, que es, de hecho, la siguiente: «Un sistema de $n$ ecuaciones con $n$ incógnitas determina *en general* las incógnitas». El término «en general» puede interpretarse de varias formas. Lo importante es una interpretación práctica algo vaga y tosca; un enunciado significa «en general» si vale «para la gran mayoría de casos en que es verosímil que algo cocurra naturalmente».

Al tratar un problema físico o geométrico de modo algebraico intentamos expresar una condición intuitivamente subyacente en nuestras ecuaciones. Queremos expresar una cláusula diferente de la condición en cada ecuación e intentamos cubrir toda la condición. Si logramos tener tantas ecuaciones como incógnitas esperamos que podremos determinar tales incógnitas. Tal esperanza es razonable. Nuestras ecuaciones «se produjeron naturalmente»; podemos esperar que estemos «en el caso general». Sin embargo, el ejemplo de la presente sección no se produjo naturalmente, sino que fue fabricado para mostrar la falta de absoluta certeza en el enunciado heurístico. Por tanto, este ejemplo no invalida en absoluto el principio heurístico subyacente.

En la vida diaria ocurre algo semejante. No nos asustamos demasiado de cosas que están fuera de lo corriente. Se pierden cartas, trenes descarrilan, y, sin embargo, no dejo de enviar cartas ni de montar en tren. Después de todo, la pérdida de cartas y las catástrofes ferroviarias son muy poco frecuentes; sólo a un pequeñísimo porcentaje de cartas o trenes les ocurren tales accidentes. ¿Por qué va a ocurrir precisamente

ahora? De modo semejante, $n$ ecuaciones con $n$ incógnitas naturalmente obtenidas pueden ser insuficientes para determinar las incógnitas. Sin embargo, esto no ocurre generalmente; ¿por qué va a ocurrir precisamente ahora?

No podemos vivir ni resolver problemas sin un poco de optimismo.

## 4. Prefiriendo la conjetura más sencilla

«Simplex sigillum veri», o «la sencillez es el sello de la verdad», decían los escolásticos. Hoy, como la humanidad es más vieja y rica con la considerable experiencia científica de los siglos intermedios, nos expresamos más cautamente; sabemos que la verdad puede ser enormemente complicada. Quizá los escolásticos no entendieron que la sencillez fuese un atributo necesario de la verdad, quizá lo que quisieron fue establecer un principio heurístico: «Lo que es sencillo tiene buena probabilidad de ser verdadero». Aún puede ser mejor, quizá, decir menos y limitarnos a aconsejar: «Intenta primero lo más sencillo».

Este consejo de sentido común incluye los movimientos heurísticos de la sección anterior, aunque un poco vagamente. Que el volumen cambie cuando cambia la forma no es sólo lo corriente, sino también el caso más sencillo. Es razonable intentar el caso más sencillo primero. Aun cuando estuviéramos obligados a volver a realizar un examen más apretado de posibilidades más complejas, el examen previo del caso más sencillo puede servir como una preparación útil.

Intentar lo sencillo primero es parte de una actitud ventajosa frente a los problemas chicos o grandes. Vamos a imaginar (en una simplificación excesiva y, sin duda, con alguna distorsión) la situación de Galileo cuando investigaba la ley de caída de los cuerpos. Si deseamos contar la Era de la ciencia moderna a partir de una fecha determinada, la fecha de esta investigación galileana se puede considerar la más apropiada.

Reconstruyamos la posición de Galileo. Con unos cuantos predecesores y unos pocos amigos que participaban de sus puntos de vista, fue, sin embargo, fuertemente atacado por la escuela filosófica dominante, la aristotélica, que ante la pregunta «¿Por qué caen los cuerpos?» se satisfacía con una explicación casi puramente verbal. Galileo preguntó: «¿Cómo caen los cuerpos?» e intentó encontrar una respuesta en la experiencia. Una respuesta precisa, expresable en números y conceptos matemáticos. Esta sustitución de «Cómo» por «Por qué», la busca de una respuesta experimental, y de una ley matemática que condensase los

hechos experimentales son un lugar común de la ciencia moderna, pero en tiempo de Galileo fueron innovaciones revolucionarias.

Una piedra que cae desde una altura choca duramente contra el suelo. Un martillo al caer desde un punto alto se introduce profundamente en la tierra. Cuanto más distancia recorre el cuerpo que cae desde su punto de partida con más velocidad se mueve —esto queda claro con una sencilla observación—. *¿Qué es lo más sencillo?* Parece bastante sencillo asumir que la velocidad de un cuerpo que cae empezando desde la posición de descanso es *proporcional a la distancia recorrida.* «Este principio parece muy natural —dijo Galileo— y corresponde a nuestra experiencia con máquinas que operan por percusión.» Galileo rechazó la proporcionalidad de la velocidad a la distancia, como «no meramente falsa, sino imposible» *.

Las objeciones de Galileo contra el hecho de asumir lo que se le presentaba como tan natural pueden ser formuladas más clara y sorprendentemente al principio en la notación del cálculo. Esto es, naturalmente, un anacronismo; el cálculo se inventó después de Galileo y, en parte al menos, bajo el impacto de sus descubrimientos. Pero usemos el cálculo. $t$ denota el tiempo pasado desde el principio de la caída y $x$ la distancia recorrida. Luego, la velocidad es $dx/dt$ (uno de los éxitos de Galileo fue formular un concepto claro de la velocidad). $g$ será una constante positiva apropiada. Esta «sencilla asunción», la proporcionalidad de la velocidad a la distancia recorrida, se expresa por la ecuación diferencial

$$\frac{dx}{dt} = gx. \tag{1}$$

Tenemos que añadir la condición inicial

$$x = 0 \quad \text{como} \quad t = 0. \tag{2}$$

De las ecuaciones (1) y (2) se sigue que

$$\frac{dx}{dt} = 0 \text{ cuando } t = 0; \tag{3}$$

esto expresa que el cuerpo que cae comienza desde la posición de descanso.

---

* Ver *Le Opere de Galileo Galilei*, edizione nazionale, vol. 8, págs. 203, 373, 383.

Obtenemos por integración de la ecuación diferencial (1) que

$$\int \frac{dx}{x} = \int g dt$$

$$\log x = gt + \log c,$$

donde $c$ es una constante positiva. Esto da por resultado

$$x = ce^{gt}, \qquad \frac{dx}{dt} = gce^{gt}.$$

De aquí obtenemos que

$$x = c > 0, \frac{dx}{dt} = gc > 0 \text{ cuanto } t = 0$$

en contradicción con (2) y (3): un movimiento que satisface a la ecuación diferencial (1) *no puede comenzar desde la posición de descanso.* Y así, la asunción que parecía tan «natural», «lo más sencillo», es contradictoria en sí misma: «no meramente falsa, sino imposible», como Galileo expresó.

Sin embargo, ¿cuál es la «próxima cosa más sencilla»? Puede ser asumir que la velocidad de un cuerpo que cae comenzando desde la posición de descanso es proporcional al tiempo transcurrido. Esta es la bien conocida ley a la que al fin llegó Galileo. En notación moderna se expresa por la ecuación

$$\frac{dx}{dt} = gt,$$

y un movimiento que satisface esta ecuación puede, ciertamente, empezar desde la posición de descanso.

## 5. Concepción del mundo

No podemos por menos de admirar el valor intelectual de Galileo, su libertad respecto del misticismo y los prejuicios filosóficos. Sin embargo, también admiramos las realizaciones de Kepler; y Kepler, un contemporáneo de Galileo, se vio envuelto profundamente en el misticismo y los prejuicios de su tiempo.

Es difícil para nosotros comprender la actitud de Kepler. El lector moderno se ve sorprendido por un título como *Un pródromo a las disertaciones cosmográficas, que contiene el* MISTERIO CÓSMICO, *sobre la admirable proporción de las órbitas celestiales y las causas propias y genuinas del número, magnitud y movimientos periódicos de los cielos, desmostrado por los cinco sólidos geométricos regulares.* El contenido es todavía más sorprendente: astronomía mezclada con teología, geometría revuelta con astrología. Pero, a pesar de lo extravagante que estos contenidos puedan parecer, la primera obra de Kepler marca el comienzo de sus grandes descubrimientos astronómicos, y da, además, una viviente y atractiva figura de su personalidad. Su sed de conocimiento es admirable, aunque resulta casi equivalente a su hambre de misterio.

Como el título de la obra dice muy correctamente, Kepler emprendió el descubrimiento de una causa o razón del número de los planetas, de sus distancias al Sol, del período de sus revoluciones. El se preguntó, en efecto: ¿Por qué hay sólo seis planetas? ¿Por qué están sus órbitas dispuestas de esta manera? Estas preguntas nos parecen hoy bastante extrañas, pero no les parecía lo mismo a sus contemporáneos *. Un día creyó que había encontrado el secreto, y apuntó en su libro de notas: «La órbita o esfera de la Tierra es la medida de todo. Circunscribámosla con un dodecaedro: la esfera que lo ciñe es Marte. Circunscribamos a Marte con un tetraedro: la esfera que lo ciñe es Júpiter. Circunscribamos a Júpiter con un cubo: la esfera que lo ciñe es Saturno. Ahora inscribamos en la Tierra un icosaedro: la esfera contenida en él es Venus. Inscribamos en Venus un octaedro: la esfera contenida en él es Mercurio. He aquí la razón del número de los planetas».

Es decir, Kepler imagina once superficies concéntricas, seis esferas que alternan con cinco sólidos regulares. La superficie primera y más exterior es una esfera y cada superficie está rodeada por la anterior. Cada esfera es asociada con un planeta: el radio de la esfera es la distancia (distancia media) del planeta al Sol. Cada sólido regular está inscrito en el anterior, que rodea a la esfera, y circunscrito por el siguiente, rodeado por la esfera.

Y Kepler añade: «Nunca encontraré palabras para expresar la delicia de este descubrimiento».

Kepler (en este aspecto un científico moderno) comparó cuidadosamente su conjetura con los hechos. El calculó una tabla que es presentada aquí con el nombre de tabla I, ligeramente modificada.

---

* Kepler rechaza la explicación de que hay seis planetas, porque 6 es el primer "número perfecto".

TABLA I

LA TEORIA DE KEPLER COMPARADA CON LAS OBSERVACIONES

| (1) *Planetas* | (2) *Observación de Copérnico* | (3) *Teoría de Kepler* | (4) *Sólidos regulares* |
|---|---|---|---|
| Saturno. | | | |
| | .635 | .577 | Cubo. |
| Júpiter. | | | |
| | .333 | .333 | Tetraedro. |
| Marte. | | | |
| | .757 | .795 | Dodecaedro. |
| Tierra. | | | |
| | .794 | .795 | Icosaedro. |
| Venus. | | | |
| | .723 | .577 | Octaedro. |
| Mercurio. | | | |

La columna (1) enumera los planetas en orden de distancias decrecientes desde el Sol; tiene seis puestos, uno más que las siguientes columnas. La columna (2) expresa la razón de las distancias de dos planetas consecutivos desde el Sol, según Copérnico; cada razón está colocada entre las líneas que contienen los nombres de los respectivos planetas; la distancia del planeta exterior es el denominador. La columna (4) enumera los cinco sólidos regulares en el orden elegido por Kepler. La columna (3) enumera la razón de los radios de las esferas inscritos y circunscritas para el sólido correspondiente. Los números de la misma línea deben estar de acuerdo. En efecto, este acuerdo es bueno en dos casos y muy malo en los tres restantes.

Ahora, Kepler (que nos recuerda de manera menos notable a los científicos modernos) empieza a amoldar su punto de vista, modificando su conjetura original. (La principal modificación es que compara la distancia de Mercurio desde el Sol no al radio de la esfera inscrita en el octaedro, sino al del círculo inscrito en el cuadrado en que un cierto plano de simetría intersecta el octaedro.) Sin embargo, no llega a ningún acuerdo definitivo entre conjetura y observación. Aun así, se aferra a su idea. La esfera es «la figura más perfecta», y próximos a ella están los cinco sólidos regulares, conocidos por Platón como las «más nobles figuras». Kepler piensa, por un momento, que la multitud innumerable de estrellas fijas puede tener algo que ver con la multitud indistinta de sólidos irregulares. Y le parece «natural» que el Sol y los planetas, las más excelentes cosas de la Creación, se relacionen de alguna manera

con las más excelentes figuras de Euclides. Este podría ser el secreto de la Creación, el «Misterio Cósmico».

A los ojos modernos la conjetura puede parecer absurda. Conocemos muchas relaciones entre hechos observables y conceptos matemáticos, pero estas relaciones son de un carácter totalmente diferente. No conocemos ninguna relación útil que tenga alguna analogía apreciable con la conjetura de Kepler. Nos parece muy extraño que Kepler pudiera creer que hay algo profundo escondido tras el número de los planetas y se pudiera preguntar a sí mismo: ¿Por qué hay precisamente seis planetas?

Podemos estar tentados de considerar la conjetura de Kepler como una ridícula aberración. Sin embargo, consideremos la posibilidad de que algunas de las teorías que se debaten hoy respetuosamente puedan ser consideradas como aberraciones en un futuro no muy lejano, si es que no completamente olvidadas. Creo que la conjetura de Kepler es altamente instructiva. Muestra con particular claridad una cuestión que merece ser tenida en cuenta: el crédito que damos a una conjetura depende de nuestra *concepción total del mundo*, de la *atmósfera científica* de nuestro tiempo.

## 6. Inagotable

El ejemplo anterior trae al primer plano un rasgo importante del razonamiento plausible. Tratemos de describirlo con cierto grado de generalidad.

Tenemos una conjetura determinada, digamos $A$. Es decir, $A$ es una proposición claramente formulada, pero que no ha sido probada. Sospechamos que $A$ es verdadera, pero no lo sabemos realmente. Tenemos cierta confianza en nuestra conjetura, la cual puede tener, aunque no necesariamente, una base articulada. Después de trabajar prolongada y aparentemente sin éxito en algún problema surge de repente esta conjetura, que puede aparecer como el único escape posible de una situación enmarañada; puede mostrarse casi como cierta, aunque no podamos decir por qué.

Después de algún tiempo pueden surgir razones más articuladas que nos hablen distintamente en favor de $A$, aunque no la prueben: razones de analogía, de inducción, de casos relativos, de experiencia general o de la inherente sencillez de $A$ misma. Tales razones, sin proporcionar una demostración estricta, pueden hacer a $A$ muy plausible.

Sin embargo, sería una advertencia para que confiásəmos en esta conjetura sin ningún argumento formulado más distintamente.

Y percibimos dicho argumento sucesivamente. Había una cuestión clara en principio, que logramos hacer resaltar de un fondo oscuro. Sin embargo, había algo más tras esta cuestión en el fondo, puesto que después logramos extraer otro argumento claro. Y así, puede haber algo más tras cada cuestión aclarada. *Quizá nuestra confianza en una conjetura no se basa nunca en fundamentos aclarados sólo;* tal confianza puede necesitar de algún modo nuestra *concepción total* como base.

Aun así, los fundamentos plausibles son importantes, y los fundamentos plausibles aclarados son especialmente importantes. En nuestro trato con la realidad observable no podemos llegar nunca a una verdad demostrativa, tenemos que apoyarnos siempre en algún fundamento plausible. En nuestro trato con cuestiones puramente matemáticas podemos llegar a una demostración estricta. Sin embargo, puede ser muy difícil llegar a ello, y la consideración de fundamentos plausibles provisionales puede darnos un apoyo temporal y conducirnos por fin al descubrimiento del argumento demostrativo definitivo.

Las razones heurísticas son importantes aunque no prueban nada. Aclararlas es también importante, aunque detrás de cada razón aclarada puede haber algo más, algún fundamento aún más oscuro y aún más importante quizá *.

Esto sugiere otra observación: si en cada caso concreto podemos aclarar sólo unos cuantos de nuestros fundamentos plausibles, y no podemos agotarlos en ningún caso concreto, ¿cómo podríamos esperar describir exhaustivamente en abstracto las clases de fundamentos plausibles?

## 7. Asunciones heurísticas usuales

Dos de nuestros ejemplos (secciones 2 y 3) nos enseñan otro punto. Recordemos brevemente una de las situaciones y toquemos una situación similar.

Al trabajar en un problema obtenemos de fuentes aparentemente distintas tantas ecuaciones como incógnitas tenemos. Debíamos saber que $n$ ecuaciones no son suficientes para determinar $n$ incógnitas: las ecuaciones pueden ser mutuamente dependientes o contradictorias. Aun así, este caso suele ser excepcional, y por ello es bastante razonable esperar que nuestras ecuaciones determinen nuestras incógnitas. Por tanto, sigamos adelante con nuestras ecuaciones, manipulémoslas

* *How to Solve It*, pág. 224.

y veamos lo que se sigue de ellas. Si hay contradicción o indeterminación, ella misma se manifestará de algún modo. Por otro lado, si llegamos a un resultado neto, nos sentiremos inclinados a gastar tiempo y esfuerzo en una demostración estricta.

Al resolver otro problema nos vemos conducidos a integrar una serie infinita término por término. Debíamos saber que una operación tal no está siempre permitida y podría darnos un resultado incorrecto. Aun así, un caso tal suele ser excepcional, y por ello es razonable esperar que nuestra serie irá bien. Por tanto, puede ser oportuno seguir adelante, ver lo que se deduce de nuestra fórmula no completamente probada y posponer inquietudes sobre una prueba completa.

Hemos tocado aquí dos *asunciones heurísticas usuales*, una sobre sistemas de ecuaciones, la otra sobre series infinitas. En todas las ramas de las matemáticas hay asunciones semejantes, y uno de los principales haberes del experto en una de esas ramas es conocer las asunciones corrientes, saber cómo utilizarlas y hasta dónde podemos confiar en ellas.

Por supuesto que no confiaremos demasiado en ninguna intuición, ya sean asunciones heurísticas usuales, ya sean nuestras propias conjeturas. Creer sin pruebas que nuestra intuición es verdadera sería una locura. Sin embargo, trabajar con la esperanza de que nuestra intuición *puede* ser cierta no deja de ser razonable. El optimismo precavido es la actitud razonable.

## EJEMPLOS Y COMENTARIOS

**1.** De un triángulo se nos da la base $a$, la altura $h$ perpendicular a $a$ y al ángulo $\alpha$ opuesto a $a$. Podríamos: ($a$) construir el triángulo, ($b$) calcular su área. ¿Son éstos todos los datos necesarios?

**2.** De un trapezoide se nos dan la altura $h$ perpendicular a los dos lados paralelos, la línea media $m$ que es paralela a los dos lados paralelos, y está a la misma distancia de ambos, y los ángulos $\alpha$ y $\beta$ entre uno de los dos lados paralelos y los dos lados oblicuos que quedan. Podríamos: ($a$) construir el trapezoide, ($b$) calcular su área. ¿Son éstos todos los datos necesarios?

**3.** Una zona es una porción de la superficie de la esfera contenida entre dos planos paralelos. La altura de la zona es la distancia de los dos planos. Dado $r$ el radio de la esfera, $h$ la altura de la zona y $d$ la distancia del plano limitante cuyo centro está más próximo al centro de la esfera, encontrar la superficie de la zona. ¿Alguna observación?

**4.** Una primera esfera tiene el radio $a$. Una segunda esfera, con

radio $b$, intersecta la primera esfera y pasa por su centro. Calcular el área de esa porción de la superficie de la segunda esfera que yace dentro de la primera esfera. ¿Alguna observación? Comprobar los casos extremos.

**5.** Reconsiderar el ejemplo de la sección 2 y probar la solución.

**6.** Un segmento esférico es una porción de la esfera contenida entre dos planos paralelos. Su superficie consta de tres partes: de una zona de la esfera y de dos círculos, llamados base y base superior del segmento. Utilizamos la siguiente notación:

$a$ es el radio de la base;

$b$, el radio de la base superior;

$h$, la altura (la distancia entre la base y la superior);

$M$, el área de la sección media (paralela a la base inferior y superior, y a la misma distancia de ambas);

$V$, el volumen del segmento.

Dados $a$, $b$ y $c$, encontrar $Mh - V$.

¿Alguna observación? Comprobar algunos casos extremos.

**7.** El eje de un cono pasa por el centro de una esfera. La superficie del cono intersecta la superficie de la esfera en dos círculos y divide la

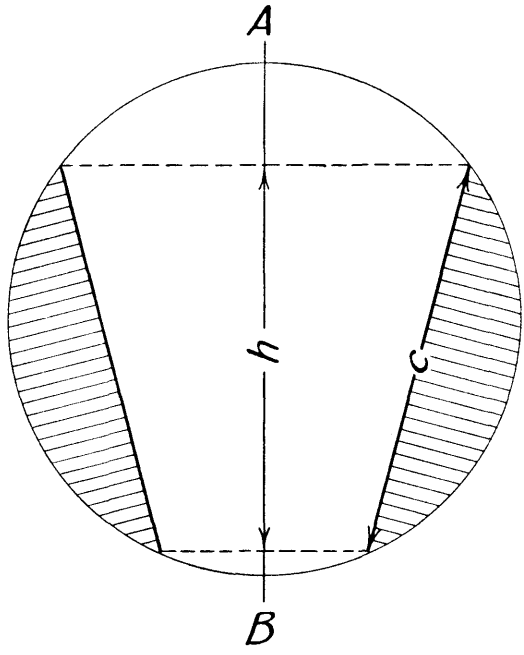



Fig. 11.2. La esfera cónicamente perforada.

esfera sólida en dos porciones: la «esfera perforada cónicamente» y la «cuña» (ver la figura 11.2, que gira sobre la línea $AB$); la cuña está dentro del cono. $r$ denota el radio de la esfera, $c$ la longitud de la cuerda que al girar engendra el agujero cónico y $h$ (la altura de la esfera perforada) la proyección de $c$ sobre el eje del cono. Dado $r$, $c$ y $h$, encontrar el volumen de la esfera cónicamente perforada. ¿Alguna observación?

**8.** El eje de un paraboloide de revolución pasa por el centro de una esfera y las dos superficies intersectan en dos círculos. Calcular el sólido de forma anular (dentro de la esfera y fuera del paraboloide) dado el radio $r$ de la esfera, $h$ la proyección del sólido de forma anular sobre el eje del paraboloide y $d$ la distancia del centro de la esfera desde el vértice del paraboloide. (Girar la figura 11.3 sobre $OX$.) ¿Alguna observación?

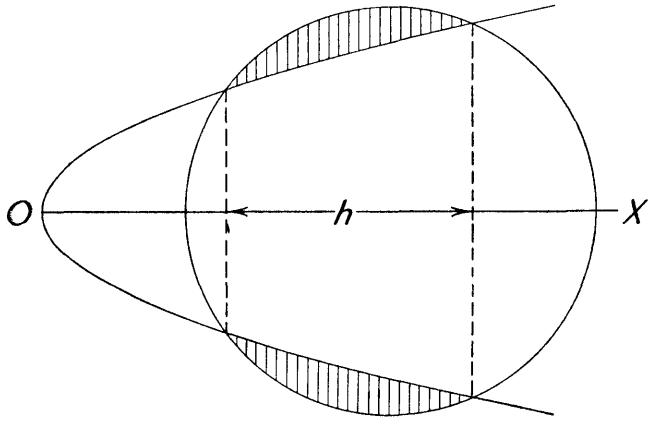



Fig. 11.3. La esfera parabólica perforada.

**9.** De un trapezoide nos dan la base inferior $a$, la base superior $b$ y la altura $h$, perpendicular a ambas bases; $a > b$. El trapezoide, girando sobre su base inferior, describe un sólido de revolución (un cilindro cubierto con dos conos); encontrar: ($a$) el volumen, y ($b$) el área de la superficie. ¿Son estos datos suficientes para determinar la incógnita?

**10.** Diez números tomados en un orden definido, $u_1, u_2, u_3, \ldots u^{10}$, están conectados de modo que desde el tercero en adelante cada uno de ellos es la suma de los dos números anteriores:

$$u_n = u_{n-1} + u_{n-2}, \quad \text{para} \quad n = 3, 4, \ldots 10.$$

Dado $u_7$, encontrar la suma de los diez números $u_1 + u_2 + \ldots + u_{10}$. ¿Son estos datos suficientes para determinar la incógnita?

**11.** Calcular $$\int_0^\infty \frac{dx}{(1+x^2)(1+x^a)}$$

¿Alguna observación? Comprobar los casos $\alpha = 0$, $\alpha \to \infty$, $\alpha \to -\infty$.

**12.** Generalizar el ejemplo 11. [Intentar primero lo más sencillo.]

**13.** Escribir una ecuación con una incógnita que no determine dicha incógnita.

**14.** Una ecuación puede determinar varias incógnitas si la naturaleza de éstas queda restringida por una condición adicional conveniente. Por ejemplo, si $x$, $y$ y $z$ son números reales completamente determinadas por la ecuación

$$x^2 + y^2 + z^2 = 0.$$

Encontrar todos los sistemas de enteros positivos $x$, $y$ satisfaciendo la ecuación $x^2 + y^2 = 128$.

**15.** Encontrar todos los sistemas de positivos enteros, $x, y, z, w$, que satisfacen la ecuación $x^2 + y^2 + z^2 + w^2 = 64$.

**16.** *El caso general.* Considerar el sistema de tres ecuaciones lineales con tres incógnitas:

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1, \\ a_2x + b_2y + c_2z &= d_2, \\ a_3x + b_3y + c_3z &= d_3. \end{aligned}$$

Asumir que los doce números dados $a_1$, $b_1$, $c_1$, $d_1$, $a_2$, ... $d_3$ son reales. El sistema se llama *determinado* si tiene una sola solución (sólo un conjunto $x$, $y$, $z$ de tres números que lo satisfacen), *indeterminado* si hay un número infinito de soluciones e *inconsistente* si no tiene solución. Visto desde varios puntos de vista, el caso en que el sistema está determinado aparece como el caso general, usual, normal, regular, y los otros aparecen como excepcionales, inusuales, anormales, irregulares.

(*a*) Geométricamente, podemos interpretar el conjunto de tres números $x$, $y$, $z$ como un punto en un sistema de coordenadas rectangulares y cada ecuación como el conjunto de puntos que lo satisfacen, como un plano. (Para esta interpretación debemos asumir, de hecho, que en el lado izquierdo de cada ecuación hay al menos un coeficiente que no desaparece, sino que debemos asumir.) El sistema de tres ecuaciones es determinado si los tres planos tienen sólo un punto común. Cuando tiene dos puntos comunes tienen una línea recta en común y el

sistema es por ello indeterminado. Cuando los tres planos son paralelos a la misma línea recta, pero no hay un punto común a los tres, el sistema es inconsistente. Si los tres planos están en una «posición general», cuando han sido «elegidos al azar», tienen sólo un punto en común y el sistema es determinado.

(*b*) Algebraicamente, el sistema de tres ecuaciones es determinado si, y sólo si, el determinante de los nueve coeficientes del lado izquierdo no desaparece. Por tanto, el sistema es determinado, a menos que una condición particular o *restricción* sea impuesta sobre los coeficientes, en forma de una ecuación.

(*c*) Podemos interpretar el conjunto de los nueve coeficientes (reales) $a_1$, $a_2$, $a_3$, $b_1$, ... $c_3$, como un punto en un espacio de nueve dimensiones. Los puntos correspondientes a sistemas que no están determinados (indeterminados o inconsistentes) satisfacen a una ecuación (con determinante $= 0$), y así forman un agregado de *dimensión más baja* (una «hipersuperficie» de ocho dimensiones).

(*d*) Es *infinitamente improbable* que un sistema de tres ecuaciones lineales con tres incógnitas dadas al azar no sea determinada. Cf. ejemplo 14.23.

**17.** Considerar la esfera inscrita y la circunscrita, y calcular la razón de los radios de estas dos esferas para cada uno de los cinco sólidos regulares.

**18.** La columna (3) de la tabla I no cambiaría si intercambiamos el cubo y el octaedro o el dodecaedro y el icosaedro. Esto dejaría la teoría de Kepler embarazosamente indeterminada. Sin embargo, Kepler muestra una singular ingenuidad descubriendo razones por las cuales uno de estos cinco nobles sólidos sería más noble y tendría precedencia sobre otro, como un barón la tiene sobre un *baronet.*

Encontrar alguna sencilla propiedad geométrica que distinga los tres sólidos que Kepler colocó alrededor de la órbita terrestre de los dos que colocó en esta órbita.

**19.** *Ninguna idea es realmente mala.* «Muchas intuiciones han resultado erróneas, y, sin embargo, útiles por conducirnos a alguna intuición mejor.» «Ninguna idea es realmente mala, a menos que no tengamos capacidad crítica. Lo que es realmente malo es carecer de ideas» *. Yo utilizo estas frases casi diariamente para consolar a algún que otro estudiante que se adelanta con alguna honesta pero ingenua idea. Estas frases se aplican tanto a las situaciones de la vida diaria

* *How to Solve It*, págs. 207-208.

como a la investigación científica. Y más espectacularmente al caso de Kepler.

Para Kepler mismo, con su mente en esa singular transición del punto de vista medieval al moderno, su idea de combinar los seis planetas con los cinco sólidos regulares apareció como radiante. Sin embargo, no puedo imaginar que Galileo, contemporáneo de Kepler, pudiera concebir tal idea. Para una mente moderna esta idea debe aparecer mala desde el principio, puesto que tiene muy poca relación con el resto de nuestro conocimiento sobre la naturaleza. Aun que está más de acuerdo con las observaciones, la conjetura de Kepler tiene un soporte muy débil, pues le falta el apoyo de la analogía.

Sin embargo, la conjetura de Kepler que resultó equivocada fue de una gran utilidad para conducirle a otra mejor. Kepler volvió a examinar las distancias medias de los planetas, sus órbitas, sus tiempos de revolución para los que esperaba encontrar alguna «explicación» semejante, hasta llegar a las célebres leyes del movimiento de los planetas, a Newton y a nuestro moderno panorama científico.

**20.** *Algunas usuales asunciones heurísticas.* Este tema merece un tratamiento más completo, pero no tenemos más remedio que limitarnos a una corta enumeración y a unos cuantos comentarios esbozados. Debemos interpretar las palabras «en general» en un sentido «práctico» y necesariamente algo vago.

«Si en un sistema de ecuaciones hay tantas ecuaciones como incógnitas, las incógnitas están determinadas, *en general.*»

Si en un problema hay tantas «condiciones» como parámetros útiles es razonable empezar con la asunción tentativa de que el problema tiene una solución. Por ejemplo, una forma cuadrática de $n$ variables tiene $n(n+1)/2$ coeficientes y una sustitución ortogonal de $n$ variables depende de $n(n-1)/2$ parámetros. Por tanto, es bastante plausible desde el principio que, por una sustitución ortogonal conveniente, cualquier forma cuadrática de $n$ variables pueda ser reducida a la expresión

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2;$$

$y_1, y_2, \ldots y_n$ son las nuevas variables introducidas por la sustitución y $\lambda_1, \lambda_2, \ldots \lambda_n$ parámetros convenientes. De hecho, esta expresión depende de $n$ parámetros y

$$n(n+1)/2 = n(n-1)/2 + n.$$

Esta observación, que viene tras una prueba de la proposición en los casos particulares $n = 2$ y $n = 3$, y una explicación del significado geo-

métrico de estos casos, puede crear una presunción bastante fuerte en favor del caso general.

«Dos operaciones límites son, *en general*, conmutativas.»

Si una de las operaciones límite es la suma de una serie infinita y la otra es la integración, nos encontramos con el caso mencionado en la sección 7 *.

«Lo que es verdadero hasta el límite es verdadero en el límite, *en general*» **.

Dado $a_n > 0$ y $\lim_{n\to\infty} a_n = a$, *no podemos* concluir que $a > 0$; meramente que $a \geqq 0$ es cierto. Consideramos una curva como el límite de un polígono inscrito y una superficie como el límite de un poliedro inscrito. Calcular la longitud de la curva, como el límite de la longitud de un polígono inscrito proporciona el resultado correcto, aunque calcular el área de la superficie como el límite del área de un poliedro inscrito puede proporcionar un resultado incorrecto ***. Aunque pueda desorientarnos fácilmente, el principio heurístico establecido es el más fértil en sugerencias e inspiraciones. Ver, por ejemplo, el ejemplo 9.24.

«Mirar una función desconocida, *en principio*, como monotónica.»

Vimos algo parecido a esto en la sección 2 cuando asumimos que con el cambio de la forma de un cuerpo su volumen cambia también, y nos sentimos desorientados. Sin embargo, el principio establecido es a menudo útil. Podemos probar una desigualdad de la forma

$$\int_a f(x)dx < \int_a g(x)dx$$

donde $a < b$. Podemos intentar probar más, es decir, que

$$f(x) < g(x).$$

Esto encoge la asunción inicial que la función con derivada $g(x)$—$f(x)$ es monotónica. (El problema es comparar los valores de esta función para $x = a$ y $x = b$.) El principio establecido está contenido en el principio heurístico más general «intentar primero lo más sencillo».

«*En general*», una función puede ser extendida en una serie de valor, cuyo primer término es una aproximación aceptable y en los demás términos tomamos lo mejor que alcanza la aproximación.»

---

* Véase G. H. HARDY, *A Course of Pure Mathematics*, 7.ª ed., págs. 493-496.

** Cf. WILLIAM WHEWELL, *The Philosophy of the Inductive Sciences*, nueva edición, volumen I, pág. 146.

*** Véase H. A. SCHWARZ, *Gesammelte Mathematische Abhandlungen*, vol. 2, páginas. 309-311.

Sin la restricción «en general» bien entendida este enunciado sería enormemente falso. Sin embargo, físicos, ingenieros y otros científicos que aplican el cálculo a su ciencia parecen ser muy aficionados a él. Esto incluye otro principio más amplio que el que establecimos previamente: «Considerar una función desconocida, *en principio*, como lineal». De hecho, si tenemos la extensión

$$f(x) = a_0 + a_1x + a_2x^2 + a_nx^3 + \ldots,$$

podemos tomar, aproximadamente,

$$f(x) \sim a_0 + a_1x.$$

(Observar que Galileo, que no conocía el cálculo, tenía ya una fuerte preferencia por la función lineal; ver la sección 4.) El presente principio está por debajo de la importancia a menudo atribuida al término inicial del error relativo; ver sección 5.2. El principio fue a menudo útil para sugerir alguna idea cercana a la verdad, aunque a veces puede alejarnos de ella.

En efecto, un físico (o un ingeniero o biólogo) pueden llegar a creer que una cantidad física $y$ depende de otra cantidad física $x$ de modo que se da la siguiente ecuación diferencial:

$$\frac{dy}{dx} = f(y).$$

Ahora bien, la integración envuelta por esta ecuación puede ser demasiado difícil, o la forma de la función $f(y)$ puede ser desconocida. En ambos casos el físico desarrolla la función $f(y)$ en valores de $y$, y puede considerar las ecuaciones diferenciales a partir de aquí como aproximaciones sucesivas:

$$\frac{dy}{x} = a_0,$$

$$\frac{dy}{dx} = a_0 + a_1 y,$$

$$\frac{dy}{dx} = a_0 + a_1y + a_2y^2.$$

Sin embargo, las curvas que satisfacen estas tres ecuaciones son de naturaleza muy diferente y la aproximación puede llegar a ser totalmente desorientadora. Afortunadamente, los físicos confían más en

cuidadosos juicios que en matemáticas cuidadosas, y de este modo obtienen buenos resultados por procedimientos semejantes, aun en los casos en que la falacia matemática es menos obvia y, por tanto, más peligrosa que en nuestro ejemplo.

**21.** *Optimismo compensatorio.* Las cantidades $a$, $b$, $c$, $d$, $e$, $f$, $g$ y $h$ son dadas. Investigar si el sistema de cuatro ecuaciones para cuatro incógnitas $x$, $y$, $u$ y $v$.

$$(S) \qquad \begin{aligned} ax + by + cv + du &= 0, \\ ex + fy + gv + hu &= 0, \\ hx + gy + fv + eu &= 0, \\ dx + cy + bv + au &= 0, \end{aligned}$$

admite alguna solución diferente de la trivial $x = y = u = v = 0$. Este sistema $(S)$ tiene, como sabemos, una solución no trivial si, y sólo si, su determinante desaparece, pero deseamos evitar el cálculo directo de esta determinante de cuatro filas. La peculiar simetría del sistema $(S)$ puede sugerir hacer

$$u = x, \qquad v = y.$$

Entonces la primera ecuación del sistema $(S)$ coincide con la cuarta, y la segunda con la tercera, de modo que el sistema de cuatro ecuaciones se reduce a un sistema de dos ecuaciones distintas

$$\begin{aligned} (a + d)x + (b + c)y &= 0, \\ (e + h)x + (f + g)y &= 0. \end{aligned}$$

Este sistema admite una solución no trivial si, y sólo si, su determinante desaparece.

Aún podemos reducir el sistema $(S)$ haciendo

$$u = -v, \quad v = -y.$$

De nuevo obtenemos dos ecuaciones distintas

$$\begin{aligned} (a - d)x + (b - c)y &= 0, \\ (e - h)x + (f - g)y &= 0. \end{aligned}$$

La desaparición de la determinante de cualquier sistema de dos ecuaciones envuelve la desaparición de la determinante del sistema $(S)$. De aquí podemos sospechar (si somos lo bastante optimistas) que esta

última determinante con cuatro filas es el *producto* de los otros dos determinantes, cada uno a su vez con dos filas.

(*a*) Probar esto y generalizar el resultado para $n$ filas.

(*b*) ¿En qué sentido hemos sido optimistas?

**22.** Tomemos el sistema de coordenadas como en la sección 9.4. El eje de las $x$ es horizontal y el de las $y$ se dirige hacia abajo. Unir el origen al punto $(a, b)$:

(1) por una línea recta,

(2) por un arco circular con centro en el eje de las $x$.

Un punto material, que se pone en movimiento desde la posición de descanso en el origen, alcanza el punto $(a, b)$ en el tiempo $T_1$ o $T_2$ a medida que se desliza hacia abajo (sin fricción) siguiendo las trayectorias (1) o (2). Galileo sugirió, como vimos en la sección 9.4, que $T_1 > T_2$. Después de algunas operaciones esta desigualdad resulta equivalente a la siguiente:

$$\int_0^h [x(1-x)]^{-3/4}\,dx < 4h^{1/4}\,(1-h)^{-1/4}$$

si establecemos

$$a^2\,(a^2+b^2)^{-1} = h.$$

Podríamos intentar probar la desigualdad extendiendo ambos lados con valores de $h$. ¿Cuál sería la posibilidad más sencilla (o la «más optimista»)?

**23.** *El cálculo numérico y el ingeniero.* El profano se inclina a pensar que los cálculos numéricos del científico son infalibles, pero oscuros. En realidad, el cálculo numérico puede ser una interesante aventura, aunque incierta. Los antiguos astrónomos intentaron, igual que lo intentan los ingenieros modernos, obtener resultados numéricos de fenómenos imperfectamente conocidos con instrumentos matemáticos tan imperfectamente conocidos como aquéllos. No puede sorprendernos que tales intentos fracasen; lo sorprendente es que a menudo tengan éxito. Aquí damos un ejemplo típico. (Los detalles técnicos, que suprimimos aquí, serán publicados en otra parte.)

Un ingeniero desea calcular cierta cantidad física $Q$ relacionada con un cuadrado de lado 1. (En efecto, $Q$ es la rigidez torsional de una viga de sección transversal cuadrada, pero el lector no necesita saber esto —de hecho, ni siquiera necesita saber lo que es rigidez torsional—.) Una solución exacta tropieza con dificultades matemáticas, y así el ingeniero recurre, como suelen hacer los ingenieros, a aproxima-

ciones. Siguiendo un cierto método de aproximación divide el cuadrado dado en «elementos» iguales, esto es, $n^2$ cuadrados más pequeños cada uno del área $1/n^2$. (Al aproximarnos a una integral doble también dividimos el área dada en elementos de esta manera.) Se puede esperar que el valor aproximado tiende al valor verdadero a medida que $n$ tiende a infinito. En efecto, no obstante, a medida que $n$ aumenta,

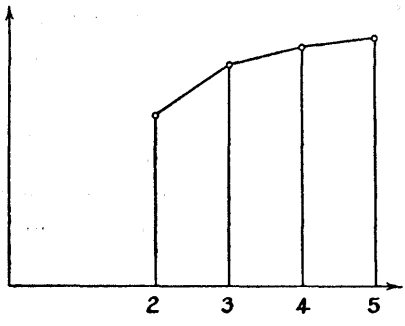



Fig. 11.4. Un ensayo: la abscisa es $n$.

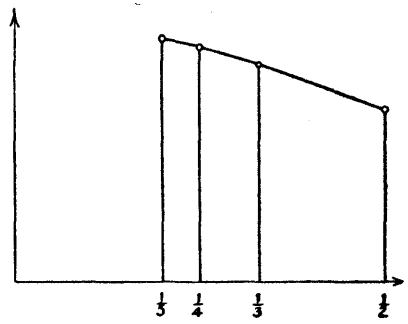



Fig. 11.5. Otro ensayo: la abscisa es $1/n$

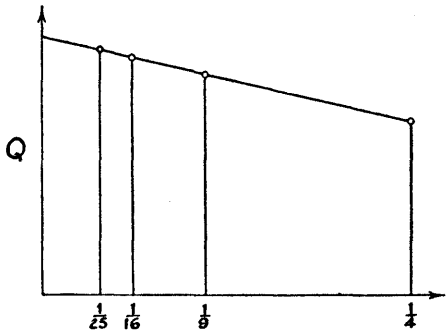



Fig. 11.6. La abscisa es $1/n^2$: ¡éxito!

también aumenta la dificultad del cálculo, y esto de forma tan rápida que pronto resulta imposible de manejar. El ingeniero considera sólo los casos $n = 2, 3, 4, 5$ y obtiene los correspondientes valores aproximados de $Q$:

0.0937 0.1185 0.1279 0.1324.

No olvidemos que estos números corresponden a los valores

1/4 1/9 1/16 1/25

de un pequeño cuadrado utilizado en cálculo, respectivamente.

El ingeniero hace un diagrama con estos resultados. Decide hacer

de los valores aproximados obtenidos para $Q$ las ordenadas, pero duda en la elección de la abscisa. Primero prueba con $n$ como abscisa, luego con $1/n$, y, finalmente, con $1/n^2$ (que es el valor numérico del área del cuadrado pequeño utilizado en la aproximación): véase figuras 11.4, 11.5, 11.6, respectivamente. La elección final es la mejor; los cuatro puntos de la figura 11.6 *están casi en la misma línea recta.* Al observar esto, el ingeniero continúa la línea hasta que ésta corta al eje vertical y considera la ordenada de la intersección como una «buena» aproximación a $Q$.

(*a*) ¿Por qué? ¿Cuál es la idea subyacente?

(*b*) Comprobar numéricamente la figura 11.6: unir cada punto al siguiente por medio de una recta y calcular las tres pendientes.

(*c*) Elegir los dos puntos más seguros de la figura 11.6, usar la línea recta que pasa por ellos en la construcción del ingeniero, calcular la aproximación a $Q$ que resulta y compararla con el verdadero valor de $Q$: 0.1406.

# Observación

El lector que ha leído los capítulos anteriores y trabajado en algunos de sus problemas tuvo una buena oportunidad para informarse de algunos aspectos del razonamiento plausible. La pretensión de los cinco capítulos restantes, recogidos en la segunda parte de esta obra, es formar una idea general de la naturaleza de dicho razonamiento. Esta pretensión tiene, según creo, considerable interés teórico e incluso puede tener también algún valor práctico; realizaremos mejor una tarea concreta cuanto mejor comprendamos la idea abstracta que la sustenta.

La formulación de ciertos patrones de razonamiento plausible es el objetivo principal de la segunda parte de esta obra. Estos patrones serán extraídos de ejemplos concretos y discutidos en estrecho contacto con ellos. Por tanto, esta última parte añadirá varios ejemplos matemáticos a los ya tratados en la primera parte y su tratamiento se hará de la misma manera.

# SEGUNDA PARTE

# PATRONES DE INFERENCIA PLAUSIBLE

# Algunos patrones notables

*En esta etapa no deseo examinar la justificación lógica de esta forma de argumentación; por el momento la considero como una práctica que se puede observar en los hábitos de los hombres y de los animales.*—BERTRAND RUSSELL *.

## 1. VERIFICACIÓN DE UNA CONSECUENCIA

En la primera parte de esta obra sobre *Inducción y analogía en matemáticas* tuvimos alguna oportunidad de familiarizarnos con la práctica del razonamiento plausible. En la presente tratamos de describir esta práctica en términos generales. Los ejemplos de la primera parte han indicado ya ciertas formas o *patrones* de razonamiento plausible. En el presente capítulo intentaremos formular algunos de tales patrones explícitamente **.

Empezamos con un patrón de inferencia plausible, que es de uso tan general que podríamos extraerle de casi cualquier ejemplo. No obstante, vamos a tomar un ejemplo que aún no hemos examinado antes.

La siguiente conjetura se debe a Euler ***: *Cualquier número entero de la forma* $8n + 3$ *es la suma de un cuadrado y del doble de un número primo.* Euler no puede probar esta conjetura, y la dificultad de hacerlo parece aún mayor hoy que en los tiempos de Euler. No obstante, Euler

---

* *Filosofía*. W. W. Norton & Co., 1927, pág. 80.

** Partes de este capítulo fueron utilizadas en mi discurso «Sobre razonamiento plausible», publicado en las *Actas del Congreso Internacional de Matemáticas*, 1950, volumen 1, págs. 739-747.

*** *Opera Omnia*, serie 1, vol. 4, págs. 20-124. En este contexto Euler considera el 1 como primo; esto es necesario para explicar el caso $3 = 1 + 2 \times 1$.

verificó su aserción para todos los enteros de la forma $8n + 3$ que están por debajo de 200; para $n = 1, 2, \ldots, 10$ ver la tabla I.

$$
\begin{aligned}
11 &= 1 + 2 \times 5 \\
19 &= 9 + 2 \times 5 \\
27 &= 1 + 2 \times 13 \\
35 &= 1 + 2 \times 17 = 9 + 2 \times 13 = 25 + 2 \times 5 \\
43 &= 9 + 2 \times 17 \\
51 &= 25 + 2 \times 13 \\
59 &= 1 + 2 \times 29 = 25 + 2 \times 17 = 49 + 2 \times 5 \\
67 &= 9 + 2 \times 29 \\
75 &= 1 + 2 \times 37 = 49 + 2 \times 13 \\
83 &= 1 + 2 \times 41 = 9 + 2 \times 37 = 25 + 2 \times 29 = 49 + 2 \times 17
\end{aligned}
$$

Tal trabajo empírico puede ser llevado adelante fácilmente; en los números que están por debajo de 1.000 * no se ha encontrado ninguna excepción. ¿Prueba esto la conjetura de Euler? De ninguna manera; ni siquiera la verificación hasta 1.000.000 lo probaría. Sin embargo, cada verificación hace la conjetura más digna de crédito y podemos ver en esto un patrón general.

Supongamos que $A$ denota una conjetura claramente formulada, la cual no está, por el momento, ni probada ni refutada. (Por ejemplo, $A$ puede ser la conjetura de Euler, de que para $n = 1, 2, 3, \ldots,$

$$8n + 3 = x^2 + 2p,$$

donde $x$ es un número entero y $p$ un número primo.) Supongamos que $B$ denota alguna consecuencia de $A$; también $B$ estará claramente establecida e, igualmente, ni probada ni refutada. (Por ejemplo, $B$ puede ser el primer caso particular de la conjetura de Euler, no anotado en la tabla I, el cual afirma que $91 = x^2 + 2p$.) Por el momento no sabemos si $A$ o $B$ son verdaderas. Sabemos, no obstante, que

$A$ implica $B$.

Ahora, intentamos la comprobación de $B$. (Unas pruebas para averiguar si desechar la afirmación sobre 91 es verdad o no.) Si decimos que $B$ es falsa podremos concluir que $A$ es falsa también. Esto es, evidentemente, cierto. Tomamos ahora un patrón de razonamiento, elemental clásico, el *modus tollens*, llamado silogismo hipotético:

$A$ implica $B$
$B$ es falsa
---
$A$ es falsa.

* Comunicación del profesor D. H. Lehmer.

La línea horizontal que separa las dos premisas de la conclusión vale por la expresión usual «luego». Tenemos aquí *inferencia demostrativa* de un tipo bien conocido.

¿Qué ocurre si $B$ resulta ser verdadera? (Realmente, $91 = 9 + 2 \times \times 41 = 81 + 2 \times 5$.) No hay conclusión demostrativa: la verificación de su consecuencia $B$ no prueba la conjetura $A$. Sin embargo, tal verificación hace a $A$ más digna de crédito. (La conjetura de Euler, verificada en un caso más, se hace de alguna manera más digna de crédito.) Tenemos aquí un patrón de *inferencia plausible:*

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ es verdadera}\end{array}}{A \text{ es más digna de crédito.}}$$

La línea horizontal vale de nuevo por «luego». A este patrón le llamaremos el *patrón fundamental inductivo* o, más brevemente, el «patrón inductivo». Este patrón inductivo no dice nada sorprendente. Por el contrario, expresa una creencia de la que ninguna persona razonable parece dudar: *La verificación de una consecuencia hace más digna de crédito a una conjetura.* Con un poco de atención podemos observar razonamientos sin cuento en la vida diaria, en los tribunales de justicia, en la ciencia, etc., que están de acuerdo con nuestro patrón.

## 2. Verificación sucesiva de varias consecuencias

En la sección presente utilizo la frase «discusión de un teorema» en el sentido específico de «discusión o examen de algunos casos particulares y de algunas consecuencias más inmediatas del teorema». Creo que la discusión de los teoremas presentados es útil tanto en las clases avanzadas como en las elementales. Vamos a considerar un ejemplo muy elemental. Supongamos que usted da una clase de geometría sólida y tiene que derivar la fórmula del área de la superficie lateral de un cono truncado. Naturalmente, el cono es recto y circular, y le dan el radio de la base $R$, el radio del círculo superior $r$, y la altura $h$. Usted hace la derivación como es usual y llega al resultado:

*A. El área de la superficie lateral del cono truncado es*

$$\pi(R + r)\sqrt{(R - r)^2 + h^2}.$$

A este teorema la llamaremos $A$ para futuras referencias.

Ahora viene la discusión del teorema $A$. Usted pregunta a la clase: ¿Pueden ustedes *comprobar le resultado?* Si no hay respuesta, usted da indicaciones más precisas: ¿Pueden ustedes comprobar el resultado *aplicándolo?* ¿Pueden ustedes comprobarlo aplicándolo a algún *caso particular que ya conozcan?* Por fin, con más o menos colaboración por parte de su clase, llegará a varios casos conocidos. Si $R = r$ obtiene usted un primer caso particular notable.

$B_1$. *El área de la superficie lateral de un cilindro es* $2\pi rh$.

Naturalmente, $h$ es la altura del cilindro y $r$ el radio de su base. Llamaremos $B_1$ a esta consecuencia de $A$ para futuras referencias. La consecuencia $B_1$ ha sido tratada ya en su clase y así sirve como una confirmación de $A$.

Otro caso particular de $A$ lo obtiene usted haciendo a $r = 0$, lo cual da por resultado:

$B_2$. *El área de la superficie lateral de un cono es* $\pi R\sqrt{R^2 + h^2}$.

Aquí $h$ es la altura del cono y $R$ el radio de su base. También esta consecuencia $B_2$ de $A$ era conocida antes, y sirve como una confirmación más de $A$.

Hay un caso particular menos obvio, pero interesante, que corresponde a $h = 0$:

$B_3$. *El área del anillo que hay entre dos círculos concéntricos con radios* $R$ *y* $r$ *es* $\pi R^2 - \pi r^2$.

Esta consecuencia $B_3$ de $A$ se estudia en geometría plana y da por resultado una nueva confirmación de $A$.

Los tres casos particulares precedentes, todos conocidos en estudios previos, apoyan a $A$ desde tres lados diferentes; las tres figuras (cilindro, cono y corona, correspondientes a $r = R$, $r = 0$, y $h = 0$, respectivamente) parecen bastante diferentes. También puede usted mencionar el caso particularísimo $r = h = 0$.

$B_4$. *El área de un círculo de radio* $R$ *es* $\pi R^2$.

A veces he observado que un muchacho de la última fila, que parecía dormir profundamente, hacia el final de mi cuidadosa derivación abrió los ojos y mostró algún interés en el progreso de la discusión. La derivación de la fórmula, aparentemente sencilla y fácil, le pareció abstrusa y difícil. No estaba convencido de la derivación. Le convence más la discusión: una fórmula que se prueba en tantos y tan diferentes casos tiene bastantes probabilidades de ser correcta, piensa. Y pensando así, concuerda con un patrón de razonamiento plausible, el cual está estre-

chamente relacionado, pero es más sofisticado que el patrón inductivo fundamental:

$A$ implica $B_{n+1}$
$B_{n+1}$ es muy diferente de las consecuencias primeramente verificadas $B_1, B_2, \ldots B_n$ *de* $A$
$B_{n+1}$ es verdadera
---
$A$ es mucho más digna de crédito.

Este patrón añade una cualificación al patrón inductivo fundamental. Ciertamente, la verificación de cualquier consecuencia fortalece nuestra creencia en una conjetura. No obstante, la verificación de ciertas consecuencias fortalece nuestra creencia más que la verificación de otras. El patrón dado trae a nuestra atención una circunstancia, que tiene una gran influencia sobre la fuerza de la evidencia inductiva: la variedad de las consecuencias probadas. La verificación de una nueva consecuencia cuenta más si la nueva consecuencia difiere más de las consecuencias primeramente verificadas.

Miremos ahora el reverso de la medalla. Tomemos el ejemplo de la anterior sección 1. Los casos sucesivos de la tabla I en que la conjetura de Euler se verifica parecen muy semejantes unos de otros —a menos que encontremos algún rasgo escondido, y parece muy difícil encontrarlo—. Además, más pronto o más tarde, nos cansaremos de esta monótona secuencia de verificaciones. Habiendo verificado un cierto número de casos, dudamos. ¿Vale la pena acometer un nuevo caso? El caso siguiente, si el resultado es negativo, podría desacreditar nuestra conjetura, pero dicho caso es tan semejante, en todos los aspectos conocidos, a los casos ya verificados, que apenas esperamos un resultado negativo. El caso siguiente, si el resultado es positivo, aumentaría nuestra confianza en la conjetura de Euler, pero este aumento de confianza es tan pequeño que escasamente vale la pena probar ese caso siguiente.

Esta consideración sugiere el siguiente patrón, el cual no es esencialmente diferente del que hemos establecido, sino que es más bien una forma complementaria suya.

$A$ implica $B_{n+1}$
$B_{n+1}$ es muy semejante a las consecuencias primeramente verificadas de $A$, $B_1, B_2, \ldots B_n$
$B_{n+1}$ es verdadera
---
$A$ es sólo un poco más digna de crédito.

***La verificación de una nueva consecuencia cuenta más o menos, según que dicha la nueva consecuencia difiera más o menos de las consecuencias primeramente verificadas.***

### 3. Verificación de una consecuencia improbable

En una breve nota*, pococo nocida, Euler considera, para valores positivos del parámetro $n$, la serie

$$(1)\quad 1 - \frac{x^2}{n(n+1)} + \frac{x^4}{n(n+1)(n+2)(n+3)} - \frac{x^6}{n\ldots(n+5)} + \ldots$$

que converge para todos los valores de $x$. Observa la suma de la serie y sus ceros para $n = 1, 2, 3, 4$.

| | |
|---|---|
| $n = 1$: suma $\cos x$, | ceros $\pm \pi/2, \pm 3\pi/2, \pm 5\pi/2, \ldots$ |
| $n = 2$: suma $(\text{sen } x)/x$, | ceros $\pm \pi, \pm 2\pi, \pm 3\pi, \ldots$ |
| $n = 3$: suma $2(1 - \cos x)/x^2$, | ceros $\pm 2\pi, \pm 4\pi, \pm 6\pi, \ldots$ |
| $n = 4$: suma $6(x - \text{sen } x)/x^2$, | no hay ceros reales. |

Euler observa una diferencia: en los primeros tres casos todos los ceros son reales, en el último caso ninguno de los ceros es real. Euler nota una diferencia más sutil entre los primeros dos casos y el tercero: para $n = 1$ y $n = 2$ la distancia entre dos ceros consecutivos es $\pi$ (a condición de que no consideremos los ceros próximos al origen en el caso $n = 2$), mas para $n = 3$ la distancia entre ceros consecutivos es $2\pi$ (con la misma condición). Esto le lleva a una sorprendente observación: en el caso $n = 3$ todos los ceros son dobles ceros. «No obstante, en Análisis estudiamos —dice Euler— que dos raíces de una ecuación siempre coinciden en la transición de raíces reales a imaginarias. Así podemos comprender por qué todos los ceros se hacen repentinamente imaginarios cuando damos a $n$ un valor que excede a 3.» Sobre la base de estas observaciones establece esta conjetura: la función definida por la *serie* (1) *tiene sólo ceros reales, y una infinidad de ellos, cuando* $0 < n \leqq 3$, *pero no tiene ceros reales en absoluto cuando* $n > 3$. En esta aseveración considera a $n$ como un parámetro continuamente variable.

En la época de Euler las cuestiones sobre la realidad de los ceros de las ecuaciones trascendentales eran totalmente nuevas, y debemos

* *Opera Omnia*, serie 1, vol. 16, sec. 1, págs. 241-265.

confesar que aún hoy no tenemos un método sistemático para decidir tales cuestiones. (Por ejemplo, no podemos probar o refutar la famosa hipótesis de Riemann.) Por lo demás, la conjetura de Euler nos parece extremadamente osada. Creo que el valor y claridad con que establece su conjetura es admirable.

No obstante, la admirable labor de Euler es comprensible hasta una cierta medida. Otros expertos realizan tareas semejantes tratando de otros temas, y cada uno de nosotros realiza algo semejante en la vida diaria. En efecto, Euler *intuyó el total a partir de unos pocos detalles desparramados*. De modo semejante, un arqueólogo puede reconstruir con bastante certeza toda una inscripción teniendo unas cuantas letras desparramadas en una piedra desgastada. Un paleontólogo puede describir, de modo que merezca confianza, todo un animal después de haber examinado algunos de sus huesos petrificados. Cuando una persona bien conocida de usted comienza a hablar de cierta manera, usted puede predecir después de unas pocas palabras todo lo que le va a contar. De la misma manera, Euler intuyó toda la situación matemática a partir de unos pocos puntos claramente reconocidos.

Aún es notable que lo intuyera a partir de tan pocos puntos, considerando sólo cuatro casos, $n = 1, 2, 3, 4$. No olvidemos, no obstante, que la evidencia circunstancial puede ser muy fuerte. Un procesado es acusado de haber hecho estallar el yate del padre de su novia; en las correspondientes pesquisas se encuenta un recibo firmado por el acusado, reconociendo la compra de cierta cantidad de dinamita. Tal evidencia fortalece enormemente la acusación. ¿Por qué? Porque la compra de dinamita en un cuidadano ordinario es un hecho bastante desusado en sí mismo, pero tal compra es completamente comprensible si el comprador intenta hacer estallar algo a alguien. Obsérvese que este caso judicial es muy semejante al caso $n = 3$ de la serie de Euler. Que todas las raíces de una ecuación escrita al azar resulten ser raíces dobles es un hecho muy desusado en sí mismo. Con todo, es completamente comprensible que en la transición de dos raíces reales a dos raíces imaginarias aparezca una doble raíz. El caso $n = 3$ es la pieza más fuerte de evidencia circunstancial producida por Euler y podemos percibir en ella un patrón general de inferencia plausible:

$A$ implica $B$

$B$ es muy improbable en sí misma

$B$ es verdadera

---

$A$ es muchísimo más digna de crédito.

También este patrón aparece como una modificación o sofisticación del patrón inductivo fundamental (sección 1). Añadamos, sin ilustración específica por el momento, el patrón complementario que explica la misma idea desde el lado inverso:

$A$ implica $B$

$B$ es bastante probable en sí misma

$B$ es verdadera

---

$A$ es sólo un poco más digna de crédito.

***La verificación de una consecuencia cuenta más o menos según sea más o menos improbable en sí misma.*** La verificación de las más sorprendentes consecuencias es muy convincente.

A propósito, Euler tenía razón: ciento cincuenta años después su conjetura ha sido completamente probada *.

## 4. Inferencia por analogía

En este punto puede ser instructivo volver la vista a los ejemplos de la primera parte sobre *inducción y analogía*. Hemos formulado unos cuantos patrones de inferencia plausible en las secciones precedentes de este capítulo: ¿cómo hacer aparecer aquellos ejemplos a la luz de estos patrones?

Volvamos a considerar dos ejemplos relacionados (de las secciones 10.1 y 10.4 del volumen I, respectivamente). Uno de estos ejemplos se relaciona con el teorema isoperimétrico y Descartes; el otro, con un teorema físico análogo del teorema isoperimétrico y lord Rayleigh. Reproducimos dos tablas del capítulo X (llamadas allí tabla I y tabla II, y aquí tabla II y tabla III, respectivamente) poniéndolas juntas. La tabla II (según la numeramos en el presente capítulo) contiene los perímetros de diez figuras, cada una de los cuales tiene la misma área 1, y en la tabla III están anotadas las principales frecuencias de las mismas diez figuras (consideradas como membranas vibrantes).

Tanto los perímetros de una tabla como las frecuencias de la otra están ordenados en sentido creciente. Las dos tablas empiezan con el círculo, que tiene el perímetro más corto entre las diez figuras puestas en lista y también la frecuencia principal más baja: esto sugiere dos teoremas.

---

* Ver el artículo del autor: «Sopra una equazione trascendente trattata da Eulero», *Bolletino dell'Unione Matematica Italiana*, vol. 5, 1926, págs. 64-68.

TABLA II

PERIMETROS DE FIGURAS DE IGUAL AREA

| | |
|---|---|
| Círculo | 3,55 |
| Cuadrado | 4,00 |
| Cuadrante | 4,03 |
| Rectángulo 3 : 2 | 4,08 |
| Semicírculo | 4,10 |
| Sextante | 4,21 |
| Rectángulo 2 : 1 | 4,24 |
| Triángulo equilátero | 4,56 |
| Rectángulo 3 : 1 | 4,64 |
| Triángulo isósceles rectángulo | 4,84 |

TABLA III

FRECUENCIAS PRINCIPALES DE MEMBRANAS DE IGUAL AREA

| | |
|---|---|
| Círculo | 4,261 |
| Cuadrado | 4,443 |
| Cuadrante | 4,551 |
| Sextante | 4,616 |
| Rectángulo 3 : 2 | 4,624 |
| Triángulo equilátero | 4,774 |
| Semicírculo | 4,803 |
| Rectángulo 2 : 1 | 4,967 |
| Triángulo isósceles rectángulo | 4,967 |
| Rectángulo 3 : 1 | 5,736 |

*De todas las figuras planas con un área dada el círculo tiene el perímetro más corto.*

*De todas las membranas con un área dada el círculo tiene la frecuencia principal más baja.*

La primera aserción es el teorema isoperimétrico; la segunda, una celebrada conjetura de lord Rayleigh. Nuestras tablas producen una sólida evidencia inductiva para ambas, pero, naturalmente, ninguna la prueba.

La situación ha cambiado desde que consideramos estas tablas en las secciones 10.1 y 10.4. Mientras tanto, hemos visto una prueba del teorema isoperimétrico (secciones 10.6-10.8, ejemplos 10.1-10.15). La propiedad mínima geométrica del círculo, apoyada inductivamente por la tabla II, ha sido probada. Es natural esperar que la propiedad mínima, física, del círculo, apoyada inductivamente por la tabla III, resulte verdadera. Esperando así, seguimos un importante patrón de inferencia plausible:

$A$ es análoga a $B$

$B$ es verdadera

---

$A$ es más digna de crédito.

*Una conjetura resulta más digna de crédito cuando una conjetura análoga resulta verdadera.*

La aplicación de este patrón a la situación discutida parece sensata. No obstante, hay nuevas indicaciones prometedoras en esta situación.

## 5. Profundización de la analogía

Las tablas II y III, puestas una junto a la otra, parecen ofrecer nuevas sugestiones. Las diez figuras consideradas no aparecen exactamente en la misma secuencia en ambas tablas. Hay algo peculiar en esta secuencia. El orden de la tabla II no parece muy diferente del seguido en la tabla III, pero ésta no es la cuestión principal. Las tablas contienen varias clases de figuras: rectángulos, triángulos, sectores. ¿Cómo están ordenadas las *figuras de la misma clase?* ¿Qué aspecto tendría una tabla más breve con figuras de una sola clase? Las tablas contienen unas cuantas figuras regulares: el triángulo equilátero, el cuadrado y, no lo olvidemos, el círculo. ¿Cómo están ordenadas las figuras regulares? ¿Podremos comparar de alguna manera figuras de diferentes clases; por ejemplo, triángulos y sectores? ¿Podríamos ampliar las bases inductivas añadiendo nuevas figuras a nuestras tablas? (Aquí estamos más restringidos. No es difícil calcular áreas y perímetros, pero la frecuencia principal es difícil de manejar y su expresión explícita es conocida sólo en muy pocos casos.) Por fin, obtendremos la tabla IV.

La tabla IV muestra un paralelismo notable entre las dos cantidades que dependen de la forma de una figura plana variable. El perímetro y la frecuencia principal. (No olvidemos que el área de la figura variable es $= 1$.) Si conocemos el perímetro no podemos de ninguna manera calcular la frecuencia principal, o viceversa. No obstante, juzgando por la tabla IV, podríamos pensar que, en muchos casos simples, estas dos cantidades *varían en la misma dirección.* Consideremos las dos columnas de datos numéricos de esta tabla y pasemos de una columna a la siguiente: si hay un incremento en una de las columnas, hay un incremento correspondiente en la otra, y si hay una disminución en una de las columnas, hay una disminución en la otra.

Vamos a centrar nuestra atención en los rectángulos. Si la razón de la longitud a la anchura incrementa de 1 a $\infty$, de manera que la forma varía desde un cuadrado hasta un rectángulo infinitamente alargado, el perímetro y la frecuencia principal parecen aumentar de modo constante. El cuadrado que, siendo una figura regular, está «más próximo» al círculo que todos los cuadriláteros, tiene el mínimo perímetro y también la mínima frecuencia principal. De los tres triángulos anotados, el triángulo equilátero, que, siendo una figura regular, está «más próximo» al círculo que todos los triángulos, tiene el mínimo perímetro y también la mínima frecuencia principal. La conducta de los sectores

TABLA IV

PERIMETROS Y FRECUENCIAS PRINCIPALES DE FIGURAS DE IGUAL AREA

| FIGURA | Perímetro | Frecuencia principal |
|---|---|---|
| Rectángulos: | | |
| 1 : 1 (cuadrado)..... | 4,00 | 4,443 |
| 3 : 2..... | 4,08 | 4,624 |
| 2 : 1..... | 4,24 | 4,967 |
| 3 : 1..... | 4,64 | 5,736 |
| Triángulos: | | |
| 60° 60° 60°..... | 4,56 | 4,774 |
| 45° 45° 90°..... | 4,84 | 4,967 |
| 30° 60° 90°..... | 5,08 | 5,157 |
| Sectores: | | |
| 180° (semicírculo) ..... | 4,10 | 4,803 |
| 90° (cuadrante)..... | 4,03 | 4,551 |
| 60° (sextante)..... | 4,21 | 4,616 |
| 45°..... | 4,44 | 4,755 |
| 36°..... | 4,68 | 4,916 |
| 30°..... | 4,93 | 5,084 |
| Figuras regulares: | | |
| círculo ..... | 3,55 | 4,261 |
| cuadrado..... | 4,00 | 4,443 |
| triángulo equilátero..... | 4,56 | 4,774 |
| Triángulos versus sectores: | | |
| tr. 60° 60° 60°..... | 4,56 | 4,774 |
| sector 60°..... | 4,21 | 4,616 |
| tr. 45° 45° 90°..... | 4,84 | 4,967 |
| sector 45°..... | 4,44 | 4,755 |
| tr. 30° 60° 90°..... | 5,08 | 5,157 |
| sector 30°..... | 4,93 | 5,084 |

es más compleja. A medida que el ángulo del sector varía desde 180° a 0°, el perímetro decrece, en primer lugar, alcanza un mínimo, y luego aumenta; y la frecuencia principal varía de la misma manera. Observemos ahora las figuras regulares. El triángulo equilátero tiene tres ejes de simetría, el cuadrado, cuatro ejes y el círculo una infinidad. Hasta donde la tabla IV nos permite observar, tanto el perímetro como la frecuencia principal parecen decrecer a medida que el número de los ejes de simetría aumenta. En la última sección de la tabla IV ponemos en contraste cada triángulo con el sector cuyo ángulo es igual al ángulo menor del triángulo. En los tres casos el sector resulta ser

«más circular», y tener el más corto perímetro y la frecuencia principal más baja.

Lo que sabemos de estas regularidades sólo llega, en definitiva, hasta donde llega la tabla IV. Esta tabla sugiere y hace plausible que estas regularidades se mantengan más allá de los límites del material experimental recogido, pero no lo prueba, de ninguna manera. Y así, la tabla IV nos lleva a varias conjeturas nuevas, que son semejantes a la conjetura de Rayleigh, aunque, naturalmente, de alcance mucho más limitado.

¿De qué manera influye la tabla IV en nuestra confianza en la conjetura de Rayleigh? ¿Podemos encontrar en la tabla IV algún fundamento razonable de dicha conjetura que no hubiéramos observado al discutir las tablas II y III?

Ciertamente, podemos encontrarlo. En primer lugar, la tabla IV contiene algunos casos particulares más en que la conjetura de Rayleigh se verifica (el triángulo de 30°, 60° y 90°, los sectores con abertura de 45°, 36° y 30°). Pero hay algo más. La analogía entre el teorema isoperimétrico y la conjetura de Rayleigh ha sido profundizada considerablemente; los hechos anotados en la tabla IV añaden varios aspectos nuevos a esta analogía. Ahora parece ser razonable considerar una conclusión de la analogía como fortaleciéndose si la analogía misma, sobre la que se basa la conclusión, se hace más fuerte. Y así, la tabla IV refuerza considerablemente el caso de Rayleigh.

## 6. Inferencia analógica matizada

No obstante, aún hay más. Como hemos observado, la tabla IV sugiere varias conjeturas que son análogas a la conjetura de Rayleigh, pero de alcance más limitado que ella. Esta tabla sugiere estas conjeturas y las presta también alguna plausibilidad. Sin embargo, esta circunstancia suscita de alguna manera, y bastante razonablemente, la plausibilidad de la conjetura original de Rayleigh. Si usted piensa así también, está de acuerdo con el siguiente patrón:

$A$ es análoga a $B$
$B$ es más digna de crédito
---
$A$ es algo más digna de crédito.

*Una conjetura se hace algo más digna de crédito cuando se hace más digna de crédito una conjetura análoga.* Esta es una forma debilitada o *matizada* del patrón formulado en la sección 4.

## EJEMPLOS Y COMENTARIOS

**1.** La tabla I, que muestra alguna evidencia inductiva para la conjetura de Euler mencionada en la sección 1, es muy semejante a la tabla de la sección 1.3, o a las tablas I, II y III del capítulo IV, o a la tabla de Euler, dada en apoyo de su «Muy Extraordinaria Ley de los números concerniente a la suma de sus divisores»; ver sección 6.2. Estas tablas se asemejan también a dos tablas dadas en el capítulo III, una en la sección 3.1 (poliedros), la otra en la sección 3.12 (particiones del espacio). ¿A cuál de estas dos es mayor la semejanza?

**2.** Habiendo verificado Euler su «Muy Extraordinaria Ley» (cf. sección 6.2) para $n = 1, 2, 3, 4, \ldots 20$, procedió a verificarla para $n = 101$ y $n = 301$. El tuvo buena razón para examinar el 101 y el 301 más que el 21 y el 22 (los cuales establece claramente al principio del número 7 de su memoria). Ignorando o recordando sólo vagamente el contenido de la ley de Euler, ¿pensaría usted que la verificación de sus dos casos (101 y 301) tiene más valor probativo que la verificación de los dos casos siguientes (21 y 22)?

**3.** Supongamos que los lados de un triángulo son $a$, $b$, $c$; el perímetro $2p = a + b + c$, y el área $A$.

Compruebe la fórmula de Herón

$$A^2 = p(p - a)\ (p - b)\ (p - c)$$

de tantas maneras como pueda.

**4.** Consideremos un cuadrilátero inscrito en un círculo. Sean los lados $a$, $b$, $c$ y $d$, el perímetro $2p = a + b + c + d$, y $A$ el área.

Se afirma que

$$A^2 = (p - a)(p - b)(p - c)(p - d).$$

Compruebe esta afirmación de tantas maneras como sea posible. ¿Tiene usted algo que comentar sobre esto?

**5.** El volumen de un tetraedro es $V$ y

$$a,\ b,\ c,$$
$$e,\ f,\ g,$$

las longitudes de sus seis aristas; las aristas $a$, $b$, y $c$ terminan en el mismo vértice del tetraedro, $e$ es la arista opuesta a $a$, $f$ a $b$ y $g$ a $c$.

(Dos aristas de un tetraedro son llamadas opuestas si no tienen vértice en común.) Las aristas $e, f$ y $g$ son los tres lados de la cara del tetraedro, opuesta al vértice en que $a$, $b$ y $c$ terminan. Se afirma que

$$144V^2 = 4a^2b^2c^2 + (b^2 + c^2 - e^2)\ (c^2 + a^2 - f^2)\ (a^2 + b^2 - g^2) - a^2(b^2 + c^2 - e^2)^2 - b^2(c^2 + a^2 - f^2)^2 - c^2(a^2 + b^2 - g^2)^2.$$

Comprobar esta aserción de tantas maneras como sea posible. [¿Es simétrica en las seis aristas la expresión propuesta para $V$?]

**6.** Establecer

$$a^n + b^n + c^n = s_n$$

para $n = 1, 2, 3, \ldots$ y definir $p$, $q$ y $r$ por la identidad de $x$

$$(x - a)\ (x - b)\ (x - c) = x^3 - px^2 + qx - r$$

de manera que

$$p = a + b + c,$$
$$q = ab + ac + bc,$$
$$r = abc.$$

(En la terminología usual, $p$, $q$ y $r$ son las «funciones simétricas elementales» de $a$, $b$ y $c$, y $s_n$ es una «suma de valores semejantes».) Obviamente, $p = s_1$. Se afirma que, para valores arbitrarios de $a$, $b$ y $c$,

$$q = \frac{2s_1^5 - 5s_1^2s_3 + 3s_5}{5(s_1^3 - s_3)},$$

$$r = \frac{s_1^6 - 5s_1^3s_3 - 5s_3^2 + 9s_1s_5}{15(s_1^3 - s_3)}$$

a condición de que el denominador no desaparezca. Comprobar estas fórmulas en el caso particular $a = 1$, $b = 2$, $c = 3$ y en tres casos más mostrados en la tabla:

| Caso | $a$ | $b$ | $c$ |
|---|---|---|---|
| (1) | 1 | 2 | 3 |
| (2) | 1 | 2 | —3 |
| (3) | 1 | 2 | 0 |
| (4) | 1 | 2 | —2 |

Idear nuevas comprobaciones. Especialmente, intentar generalizar los casos (2), (3) y (4).

**7.** Supongamos que $A$, $B_1$, $B_2$, $B_3$ y $B_4$ tienen el mismo significado que se les dio en la sección 2. ¿Provee la verificación de $B_4$, que viene después de $B_1$, $B_2$ y $B_3$, evidencia inductiva adicional a $A$?

**8.** Recordemos la «Muy Extraordinaria Ley» y el significado de las abreviaturas $T$, $C_1$, $C_2$, $C_3$, ... $C_1^*$, $C_2^*$, $C_3^*$, ... explicado en la sección 6.3. Euler apoyó el teorema $T$, cuando aún no podía probarlo, inductivamente, verificando sus consecuencias $C_1$, $C_2$, $C_3$, ... $C_{20}$. (Quizá continuase aún adelante.) Luego descubrió que también $C_1^*$, $C_2^*$, $C_3^*$, ... son consecuencias de $T$, y verificó $C_1^*$, $C_2^*$, ... $C_{20}^*$, $C_{101}^*$, $C_{301}^*$. Gracias a estas nuevas verificaciones la confianza de Euler fue muy reforzada, pero ¿fue reforzada con razón? [Aquí se necesita más atención al detalle que en el ejemplo 2.]

**9.** Volvamos a la conjetura de Euler discutida en la sección 1; en gracia a la brevedad la llamaremos la «conjetura $E$». Anotemos concisamente el significado de esta abreviación:

$$E: \qquad 8n + 3 = x^2 + 2p.$$

La idea que llevó a Euler a su conjetura $E$ merece mención. Euler dedicó gran parte de su obra a aquellas proposiciones celebradas en la teoría de los números que Fermant estableció sin probar. Una de éstas (la llamaremos «conjetura $F$») dice que cualquier número entero es la suma de tres números trigonales. Anotemos concisamente el significado de esta abreviación.

$$F: \qquad n = \frac{x(x-1)}{2} + \frac{y(y-1)}{2} + \frac{z(z-1)}{2}.$$

Euler observó que *si* su conjetura $E$ era verdadera la conjetura $F$ de Fermat se desprendería fácilmente de ella. Esto es, Euler se contentó con que $E$ implicase a $F$. (Para detalles, ver el próximo ejemplo, número 10.) Resuelto a probar la conjetura $F$ de Fermat, Euler deseó, naturalmente, que su conjetura fuera verdadera. ¿Es ésta una mera creencia basada en sus deseos? No lo creo; las relaciones consideradas dan por resultado algún fondo débil, pero en cierta medida razonable, para creer la conjetura $E$ de Euler, de acuerdo con el siguiente esquema:

$E$ implica $F$

$F$ es digna de crédito

---

$E$ es algo más digna de crédito.

Aquí tenemos otro modelo de inferencia plausible. El lector lo comparará con el patrón fundamental inductivo.

**10.** Al probar que $E$ implica a $F$ (en la notación del ejemplo anterior) Euler utilizó un resultado más profundo que probó previamente: un número primo de la forma $4n + 1$ es una suma de dos cuadrados. (Esto fue examinado inductivamente en el ejemplo 4.4.) Dando esto por garantizado, prueba que $E$ implica realmente a F.

**11.** Después de haber concebido su conjetura, examinada en la sección 3, Euler la probó calculando los primeros ceros de su serie para unos pocos valores de $n$. (Por «primer cero» entendemos un cero cuyo valor absoluto es un mínimo. Si $x$ es un primer cero de la serie en cuestión, también $-x$ es un cero, y $x$ y $-x$ son «primeros ceros». Por tanto, $x$ es real si, y sólo si, $x^2$ es positiva.) Naturalmente, Euler tuvo que calcular estos ceros de modo aproximativo. Un método (el método de Bernoulli) que frecuentemente utilizó para tal propósito dio por resultado las siguientes secuencias de valores aproximados para el primer cero $x$ en los casos $n = 1/2, 1/3, 1/4$.

| $n = 1/2$ | $n = 1/3$ | $n = 1/4$ |
|---|---|---|
| $4x^2 \sim 3.000$ | $9x^2 \sim 4.0000$ | $16x^2 \sim 5.0000$ |
| 3.281 | 4.2424 | 5.2232 |
| 3.291 | 4.2528 | 5.2302 |
| 3.304 | 4.2532 | 5.2304 |

En los tres casos los valores aproximados parecen tender a un límite positivo de modo rápido y regular. Euler toma esto como signo de que los primeros ceros son reales y ve en ello una confirmación de su conjetura.

Vamos a realizar el esquema general de la conclusión heurística de Euler. Supongamos que $A$ significa su conjetura explicada en la sección 3, concerniente a la realidad de los ceros de su serie. $B$, el hecho de que para $n = 1/2$ el primer cero sea real. Como es obvio, $A$ implica $B$. Ahora bien, Euler no lo probó; él sólo hizo que $B$ fuese más digna de crédito. Por tanto, tenemos aquí el siguiente modelo de inferencia plausible:

$A$ implica $B$

$B$ es más digna de crédito

---

$A$ es algo más digna de crédito.

La segunda premisa es más débil que la segunda premisa del patrón inductiva fundamental. La palabra «algo» va intercalada para enfatizar

que también la conclusión es más débil que en el patrón fundamental inductivo.

**12.** Un matemático moderno puede derivar una conclusión heurística más rigurosa de los datos numéricos del ejemplo 11 que la de Euler. Se puede mostrar que, si la serie de Euler tiene sólo ceros reales, los valores aproximados, sucesivos, obtenidos por el método de Bernoulli, forman necesariamente una secuencia *creciente* *. Vamos a hacer que $A$ signifique la misma conjetura que en el ejemplo 11, pero supongamos que $B$ denota ahora otra proposición, a saber: «Para $n=1/2$, las cuatro primeras aproximaciones, obtenidas por el método de Bernoulli, forman una secuencia creciente, y lo mismo para $n = 1/3$ y $n = 1/4$». Después sabemos que ambas premisas del patrón fundamental inductivo sostienen:

$A$ implica $B$

$B$ es verdadera

y la conclusión heurística resultante es más fuerte.

Se pueden añadir dos anotaciones a lo anterior.

(1) Euler no formuló la propiedad, que acabamos de citar, del método de Daniel Bernoulli y, ciertamente, no la probó. Sin embargo, hay bastantes probabilidades de que, estando en posesión de una vasta experiencia con este método, tuviera alguna clase de conocimiento inductivo de ella. Así, aunque Euler no realizó explícitamente la inferencia del matemático moderno, la poseyó en una forma menos clara. Y, probablemente, tuvo, en su rico fondo matemático, aún otras indicaciones, que no pudo formular completamente y que aún hoy no han podido ser aclaradas.

(2) Los datos numéricos citados en el ejemplo 11 llevan al autor a sospechar el teorema general probado l. c. Este es un ejemplo pequeño, pero concreto, del uso del método inductivo en la investigación matemática.

**13.** En el capítulo IV investigamos inductivamente la suma de cuatro cuadrados impares; ver secciones 4.3-4.6, tabla I. Después atacaremos los problemas análogos que implican cuatro cuadrados arbitrarios y ocho cuadrados; ver ejemplos 4.10 y 4.23, y las tablas II y III. La primera investigación nos ayudará a reconocer la ley en los casos posteriores. ¿Aumentará también nuestra confianza en el resultado de la última el resultado de la primera?

---

* Véase el ensayo del autor: «Notas sobre series de valores», *Acta Scientiarum Mathematicarum*, v. 12 B, 1950, págs. 199-203.

**14.** *Conclusión inductiva de esfuerzos sin fruto.* Construir mediante regla y compás el lado de un cuadrado de área igual a un círculo de radio dado. Esta es la formulación estricta del famoso problema de la cuadratura del círculo, que fue concebido por los griegos. No fue olvidado en la Edad Media, aunque no podemos creer que entonces entendiera mucha gente su estricta formulación; Dante se refiere a él en la culminación teológica de *La Divina Comedia*, hacia el final del último canto. El problema tenía ya dos mil años cuando la Academia Francesa resolvió que los manuscritos que proponían la cuadratura del círculo no serían examinados. ¿Era estrecha de miras la Academia? No lo creo así; después de los esfuerzos sin fruto de miles de personas, durante miles de años, había algún fundamento para sospechar que el problema era insoluble.

Hemos sido incitados a abandonar una tarea, que resiste a esfuerzos repetidos. Desistimos sólo después de muchos y grandes esfuerzos, si estamos obstinados o profundamente interesados. Desistimos después de unos pocos ensayos iniciales, si nos cansamos fácilmente o no estamos seriamente interesados. Sin embargo, en cualquier caso hay una especie de conclusión inductiva. La conjetura en consideración es:

*A.* Es imposible realizar esta tarea.

Observamos:

*B.* No podemos realizar esta tarea aún.

Esto es, en sí mismo, *muy* inverosímil. Pero, ciertamente,

*A* implica *B*,

y así, nuestra observación a propósito de *B* hace a *A* más digna de crédito, según el patrón inductivo fundamental.

Esta imposibilidad de cuadrar el círculo, estrictamente formulada, fue probada en 1882, por Lindemann, después del trabajo básico de Hermite. Hay otros problemas semejantes, que datan de los griegos (la trisección de un ángulo y la duplicación del cubo), los cuales, después de la evidencia acumulada de esfuerzos sin fruto, han sido dejados últimamente como insolubles. Tras los esfuerzos fracasados por construir un «móvil perpetuo» los físicos formularon el «principio de la imposibilidad de un movimiento perpetuo», y este principio resultó extraordinariamente fructífero.

# Otros patrones y primeros lazos

*Cuando hemos comprendido intuitivamente algunas proposiciones simples... es útil recorrerlas con un continuo e ininterrumpido movimiento de pensamiento para meditar sobre sus relaciones mutuas y concebir distintamente varias de ellas, tantas como sea posible, simultáneamente. De esta manera nuestro conocimiento gana certeza y la capacidad de la mente aumenta notablemente.*—DESCARTES *.

## 1. EXAMINANDO UNA CONSECUENCIA

Consideremos una situación que se da frecuentemente en la investigación matemática. Deseamos decidir si una proposición matemática $A$, claramente formulada, es cierta o no. Tenemos, quizá, alguna confianza intuitiva en la verdad de $A$, pero eso no es bastante: deseamos probar o refutar $A$. Trabajamos en este problema, pero sin éxito decisivo. Tras un rato, observamos una consecuencia $B$ de $A$. Esta $B$ es una proposición matemática claramente formulada, de la cual sabemos que se sigue de $A$:

$$A \text{ implica } B.$$

Sin embargo, no sabemos si $B$ es cierta o no. Ahora bien, parece que $B$ es más accesible que $A$; por una u otra razón tenemos la impresión de que nos irá mejor con $B$ de lo que nos había ido con $A$. Por tanto, dediquémonos a examinar $B$. Trabajamos para contestar la pregunta: ¿Es $B$ cierta o falsa? Al final logramos contestarla: *¿Cómo influye esta contestación sobre nuestra confianza en A?*

Eso depende de la contestación. Si encontramos que $B$, la consecuencia de $A$, es falsa, podemos inferir con certeza que $A$ también es falsa. Sin embargo, si encontramos que $B$ es cierta, no hay aquí inferencia demostrativa: aunque su consecuencia $B$ sea cierta, $A$ podría ser falsa.

* La once de sus reglas para la dirección de la mente. Ver *Oeuvres*, editada por Adam y Tannery, vol. 10, 1908, pág. 407.

Hay una inferencia heurística: puesto que la consecuencia, $B$, es cierta, $A$ misma parece merecer más confianza. De acuerdo con la naturaleza de nuestro resultado sobre $B$, seguimos, bien un patrón demostrativo, bien uno heurístico:

| *Demostrativo* | *Heurístico* |
|---|---|
| $A$ implica $B$ | $A$ implica $B$ |
| $B$ es falsa | $B$ es cierta |
| $A$ es falsa. | $A$ es más digna de crédito. |

Encontramos ya estos patrones en la sección 12.1, donde les dimos el nombre de patrón heurístico y patrón fundamental inductivo. Nos encontraremos con patrones semejantes aunque diferentes en las secciones siguientes.

## 2. Examinando un fundamento posible

Consideremos otra situación que se da con frecuencia en la investigación matemática. Deseamos decidir si la proposición $A$, claramente formulada, es cierta o no; deseamos probar o refutar $A$. Tras algunos tanteos, nos encontramos con otra proposición $B$, de la que se sigue $A$. No sabemos si $B$ es cierta o no, pero nos resulta satisfactorio que

$A$ está implicada en $B$.

Así, si pudiéramos probar $B$, se seguiría de ella $A$; $B$ es un fundamento posible de $A$. Podemos estar cansados de $A$, o $B$ puede parecernos más prometedora que $A$; por una u otra razón, nos dedicamos a examinar $B$. Ahora, pues, nuestro objetivo es probar o refutar $B$. Finalmente, tenemos éxito. ¿Cómo influirá el resultado concerniente a $B$ sobre nuestra confianza en $A$?

Eso depende de la naturaleza de nuestro resultado. Si encontramos que $B$ es cierta, podemos concluir que $A$, que está implicada en $B$ (se sigue de $B$, es una consecuencia suya), es también cierta. Sin embargo, si encontramos que $B$ es falsa, no hay conclusión demostrativa: $A$ podría ser todavía cierta. Pero nos hemos visto obligados a descartar un posible fundamento de $A$, tenemos una oportunidad menos para probarla; nuestra esperanza para probar $A$ a partir de $B$ ha fracasado: si hay algún cambio de nuestra confianza en $A$, como consecuencia de la

refutación de $B$, sólo puede ser un cambio para peor. En resumen, de acuerdo con la naturaleza de nuestro resultado, referente a $B$, seguimos, bien un patrón demostrativo o uno heurístico:

| *Demostrativo* | *Heurístico* |
|---|---|
| $A$ implicada en $B$ | $A$ implicada en $B$ |
| $B$ cierta | $B$ falsa |
| $A$ cierta | $A$ menos digna de crédito. |

Observemos que la primera premisa es la misma en ambos patrones. Las segundas premisas son diametralmente opuestas, y las conclusiones son también opuestas, aunque no totalmente.

La inferencia demostrativa sigue un patrón clásico, el *modus ponens* del llamado silogismo hipotético. El patrón heurístico es semejante, aunque diferente, del patrón fundamental inductivo; ver sección 1 o seción 12.1. Podemos establecer la inferencia heurística con palabras: *nuestra confianza en una conjetura puede disminuir solamente cuando un fundamento posible para ella es refutado.*

## 3. Examinando una conjetura conflictiva

Consideremos una situación que no es demasiado usual en la investigación matemática, pero que se da frecuentemente en las ciencias naturales. Examinemos dos conjeturas, $A$ y $B$, incompatibles o en conflicto. Cuando decimos que $A$ está en conflicto con $B$ o

$$A \text{ es incompatible con } B$$

queremos decir que la verdad de una de las dos proposiciones $A$ y $B$ necesariamente implica la falsedad de la otra. Así, $A$ puede ser cierta o no y $B$ puede ser cierta o no; no sabemos cuál de los dos casos es cierto, pero sabemos que ambos no pueden serlo. No obstante, podrían ser ambos falsos. Un naturalista propone la conjetura $A$ para explicar un fenómeno, otro propone la conjetura $B$ para explicar el mismo fenómeno de modo diferente. Las explicaciones son incompatibles; ambos naturalistas no pueden tener razón, pero pudiera ser que ambos estuviesen equivocados.

Si una de las dos conjeturas, digamos $B$, se ha probado que es cierta, entonces la suerte de la otra está echada definitivamente: $A$ debe estar

equivocada. Si, no obstante, $B$ ha sido refutada, la suerte de $A$ no está todavía definitivamente asegurada: podría estar $A$ también equivocada. Sin embargo, innegablemente, por la refutación de una conjetura rival, incompatible con ella, $A$ puede ganar. (El naturalista que inventó $A$, sin duda pensaría así.) Y por ello tenemos de nuevo dos patrones:

| *Demostrativo* | *Heurístico* |
|---|---|
| $A$ incompatible con $B$ | $A$ incompatible con B |
| $B$ cierta | $B$ falsa |
| $A$ falsa. | $A$ más digna de crédito. |

*Nuestra confianza en una conjetura puede aumentar cuando una conjetura rival, incompatible, es refutada.*

## 4. Términos lógicos

En las tres secciones anteriores hemos visto tres pares de patrones. Cada par consta de un patrón demostrativo y de otro heurístico; los tres patrones demostrativos se relacionan unos con otros y lo mismo ocurre entre los tres patrones heurísticos. Las relaciones entre los patrones demostrativos son relaciones claras de lógica formal. En la sección siguiente intentaremos aclarar las relaciones entre los patrones heurísticos. La sección presente nos prepara para la próxima, con la discusión de unos cuantos términos sencillos de lógica formal [2].

(1) El término *proposición* puede ser tomado en un sentido más general, pero la mayor parte del tiempo será suficiente, e incluso ventajoso, considerar alguna proposición matemática, claramente formulada, de la cual *por el momento no sabemos si es verdadera o no.* (Un buen ejemplo de proposición para un lector más preparado es la celebrada «hipótesis de Riemann»; la función de Riemann tiene solamente ceros reales. A despecho de los esfuerzos de muchos y excelentes matemáticos, no sabemos si esta proposición es verdadera o falsa.) Utilizaremos las mayúsculas $A$, $B$, $C$, ... para denotar proposiciones.

(2) La *negación* de la proposición $A$ es una proposición verdadera si, y sólo si, $A$ es falsa. Expresemos la negación de $A$ por no-$A$.

[2] Tratamos aquí la lógica formal a la vieja usanza (utilizando lenguaje ordinario y evitando símbolos lógicos hasta donde sea posible), aunque con unas cuantas ideas modernas. Algunos de los símbolos lógicos más sencillos serán usados incidentalmente más tarde, en la sección 7.

(3) Las dos proposiciones «*A* es falsa» y «no-*A* es verdadera» valen exactamente lo mismo. Podemos sustituir una por otra en cualquier contexto sin cambiar el significado, la verdad o la falsedad de todo el texto. Dos proposiciones que pueden ser sustituidas una por otra son denominadas *equivalentes*. Así, la proposición «*A* es falsa» es equivalente a la proposición «no-*A* es verdadera». Será conveniente escribir esto en la forma abreviada:

«*A* falsa» eq. «no-*A* verdadera».

También es correcto decir que:

«*A* verdadera» eq. «no-*A* falsa»,

«no-*A* verdadera» eq. «*A* falsa»,

«no-*A* falsa» eq. «*A* verdadera».

(4) Decimos que las dos proposiciones *A* y *B* son *incompatibles* entre sí si ambas no pueden ser verdaderas. La proposición *A* puede ser verdadera o falsa, lo mismo que *B*; si consideramos *A* y *B* juntamente, se pueden dar cuatro casos:

| | |
|---|---|
| *A* verdadera, *B* verdadera | *A* verdadera, *A* falsa |
| *A* falsa, *B* verdadera. | *A* falsa, *B* falsa. |

Si decimos que *A* es incompatible con *B*, entendemos que el primero de estos cuatro casos (en la esquina superior, izquierda) está excluido. La incompatibilidad es siempre mutua. Por tanto,

«*A* incompatible con *B*» eq. «*B* incompatible con *A*».

(5) Decimos que *A* implica *B* (o *B* está implicada en *A*, o *B* se sigue de *A*, o *B* es una consecuencia de *A*, etc.) si *A* y no-*B* son incompatibles. Así, el concepto de implicación se caracteriza por la siguiente equivalencia:

«*A* implica *B*» eq. «*A* incompatible con no-*B*».

Saber que *A* implica *B* es importante. Por el momento no sabemos si *A* es verdadera o no, y estamos en la misma ignoracia respecto de *B*. Si, no obstante, resultase algún día que *A* es verdadera, conoceremos inmediatamente que no-*B* tiene que ser falsa y *B* debe ser verdadera. Conocemos que

«*A* incompatible con no-*B*» eq. «no-*B* incompatible con *A*».

Conocemos también que

«no-$B$ incompatible con $A$» eq. «no-$B$ implica no-$A$».

De la cadena de las tres últimas equivalencias concluimos:

«$A$ implica $B$» eq. «no-$B$ implica no-$A$».

Esta última equivalencia es bastante importante en sí misma y será esencial en la siguiente consideración.

(6) Los pocos puntos de lógica formal discutidos en esta sección nos capacitan ya para aclarar la relación entre los patrones demostrativos, que encontramos en las tres secciones anteriores.

Empecemos por el patrón demostrativo formulado en la sección 1 (el *modus tollens*):

$$\begin{array}{c} A \text{ implica } B \\ B \text{ falsa} \\ \hline A \text{ falsa.} \end{array}$$

Se comprende que este patrón puede ser aplicado en general. Vamos a aplicarlo sustituyendo no-$A$ y no-$B$ por $A$ y por $B$, respectivamente. Obtenemos

$$\begin{array}{c} \text{no-}A \text{ implica no-}B \\ \text{no-}B \text{ falsa} \\ \hline \text{no-}A \text{ falsa.} \end{array}$$

Hemos visto, no obstante, en la sección anterior, que

«no-$A$ implica no-$B$» eq. «$B$ implica $A$»,

«no-$B$ falsa» eq. «$B$ verdadera»,

«no-$A$ falsa» eq. «$A$ verdadera».

Sustituyamos las tres proposiciones equivalentes, correspondientes, aquí expuestas, por las premisas y la conclusión del último patrón considerado. Entonces tenemos:

$$\begin{array}{c} B \text{ implica } A \\ B \text{ verdadera} \\ \hline A \text{ verdadera,} \end{array}$$

que es el patrón demostrativo de la sección 2, el *modus ponens*.

Dejamos al lector derivar de modo semejante el patrón demostrativo de la sección 3 del de la sección 1.

## 5. Lazos lógicos entre patrones de inferencia plausible

La discusión anterior fue meramente preparatoria. No hemos discutido aquellos puntos clásicos de lógica demostrativa por sí mismos, sino para prepararnos al estudio de la inferencia plausible. Consideramos la derivación del *modus ponens* a partir del *modus tollens* no con la vana esperanza de presentar algo nuevo y sorprendente referente a estas materias clásicas, sino como preparación para cuestiones como la siguiente: ¿Podemos derivar por pura lógica formal el patrón heurístico de la sección 2 del patrón heurístico de la sección 1?

Es evidente que no podemos. De hecho, estos patrones contienen proposiciones tales como «*A* se hace más digna de crédito» o «*A* se hace menos digna de crédito». Aunque todo el mundo comprende lo que esto significa, el lógico formal rehusa entender estos enunciados, y tiene razón, pues en lógica formal no hay lugar para ellos; ni medios para manejarlos.

No obstante, podríamos ampliar el dominio de la lógica formal de una manera apropiada. La cuestión principal parece ser añadir la equivalencia siguiente al contenido clásico de la lógica formal: «no-*A* se hace más digna de crédito» es equivalente a «*A* se hace menos digna de crédito». Abreviadamente,

«no-*A* más digna de crédito» eq. «*A* menos digna de crédito».

Admitiendo esto obtenemos un útil instrumento. Ahora podemos adaptar el procedimiento de la sección 4 (6) a nuestra aspiración actual y proceder como sigue. Empezamos por el patrón fundamental inductivo, dado en la sección 1. Lo aplicamos a no-*A* y no-*B* en vez de aplicarlo a *A* y *B*; esto es, sustituimos no-*A* por *A* y no-*B* por *B* en dicho patrón. Luego aplicamos tres equivalencias; dos de éstas son puramente lógicas (y, a propósito, discutidas en la sección anterior) y la tercera es la nueva equivalencia esencial, introducida en la presente sección. Mediante estos pasos llegamos, finalmente, al patrón heurístico dado en la sección 2.

El lector se puede ocupar de anotar esta derivación con detalle y derivar el patrón heurístico de la sección 3, de la misma manera, a partir del patrón fundamental inductivo de la sección 1.

## 6. Inferencia matizada

Me parece que el razonamiento inductivo en el dominio matemático es más fácil de estudiar que en el dominio físico. La razón es bastante sencilla. Al hacer una pregunta matemática se puede esperar una respuesta totalmente carente de ambigüedad, un cortante Sí o No. Al preguntar a la naturaleza no se puede esperar una respuesta sin algún margen de incertidumbre. Usted predice que un eclipse lunar empezará en tal y tal momento. En la realidad se observa el principio del eclipse cuatro minutos después de lo predicho. Si la comparamos con los logros de la astronomía griega tal predicción sería pasmosamente correcta; de acuerdo con el grado de precisión moderna resulta escandalosamente incorrecta. Una discrepancia determinada entre predicción y observación puede ser interpretada como confirmación o refutación. Tal interpretación depende de alguna clase de razonamiento plausible, cuyas dificultades empiezan antes en las «situaciones físicas» que en las «situaciones matemáticas». Intentaremos reducir esta distinción a su expresión más sencilla.

Supongamos que estamos investigando una conjetura matemática mediante el examen de sus consecuencias. $A$ denota la conjetura y $B$ una de sus consecuencias, de manera que $A$ implica $B$. Llegamos a una decisión final, que concierne a $B$: bien refutamos $B$ o la probamos, y, según esto, nos enfrentamos con una de las dos situaciones siguientes:

| | |
|---|---|
| $A$ implica $B$ | $A$ implica $B$ |
| $B$ falsa | $B$ verdadera |

A éstas las llamaremos «situaciones matemáticas». Las hemos considerado repetidamente, y sabemos qué inferencia se puede obtener de cada una de ellas.

Supongamos ahora que estamos investigando una conjetura física $A$, y que probamos experimentalmente una de sus conjeturas $B$; nuestros experimentos pueden mostrar, por otra parte, que $B$ o su contraria es más difícil de creer. De acuerdo con esto, nos enfrentamos con una de las dos situaciones siguientes:

| | |
|---|---|
| $A$ implica $B$ | $A$ implica $B$ |
| $B$ es apenas digna de crédito | $B$ es casi cierta |

Llamamos a éstas «situaciones físicas». ¿Qué inferencia es razonable en estas situaciones? (El espacio en blanco bajo la línea horizontal que sugiere la palabra «luego» simboliza la cuestión abierta.)

En cada una de las cuatro situaciones consideradas tenemos dos datos o *premisas*. La primera premisa es la misma en las cuatro situaciones; toda la diferencia entre ellas estriba en la segunda premisa. Esta segunda premisa está al nivel de la pura lógica formal en las situaciones «matemáticas», pero está a un nivel mucho más vago en las situaciones «físicas». Esta diferencia me parece esencial; y puede explicar las dificultades adicionales de las situaciones físicas.

Vamos a examinar las cuatro situaciones «con un continuo movimiento interrumpido de pensamiento», como a Descartes le gustaba decir (ver el lema al principio de este capítulo). Imaginemos que nuestra confianza en *B* cambia gradualmente, varía «continuamente». Supongamos que *B* se va haciendo cada vez menos digna de crédito, luego escasamente creíble, y, finalmente, falsa. Por otro lado, imaginemos que *B* se va haciendo más digna de crédito, luego aún más creíble, prácticamente cierta, y, por fin, verdadera. Si la *fuerza de nuestra conclusión varía continuamente en la misma dirección que la fuerza de nuestra confianza en B*, hay poca duda sobre cuál será nuestra conclusión, puesto que los dos casos extremos (*B* falsa, *B* verdadera) están claros. De esta manera llegamos a los siguientes patrones:

| | |
|---|---|
| *A* implica *B* | *A* implica *B* |
| *B* menos digna de crédito | *B* más digna de crédito |
| *A* menos digna de crédito. | *A* algo más digna de crédito. |

La palabra «algo» del segundo patrón nos recuerda que la conclusión es, naturalmente, más débil que en el patrón fundamental inductivo. *Nuestra confianza en una conjetura está influida por nuestra confianza en una de sus consecuencias y varía en la misma dirección.* Llamaremos a éstos patrones matizados; el primero es un patrón demostrativo matizado, el segundo es la versión matizada del patrón fundamental inductivo. El término «matizado» se refiere al debilitamiento de la segunda premisa: «menos digna de crédito» en vez de «falsa», «más digna de crédito» en vez de «verdadera». En la sección 12.6 hemos utilizado ya el término de este significado.

Obtuvimos los patrones matizados introducidos a partir de sus casos extremos, el *modus tollens* y el patrón fundamental inductivo examinado en la sección 1, por debilitamiento de la segunda premisa. De la misma manera podemos obtener otros patrones matizados de los patrones

formulados en las secciones 2 y 3. Enunciamos uno aquí (en la sección siguiente están todos en una lista). El patrón heurístico de la sección 3 da por resultado el siguiente patrón matizado:

$$\frac{\begin{array}{c}A \text{ incompatible con } B \\ B \text{ menos digna de crédito}\end{array}}{A \text{ algo más digna de crédito.}}$$

## 7. Una tabla

Para anotar concisamente los patrones discutidos en este capítulo será conveniente usar algunas abreviaciones. Escribimos

$$A \rightarrow B \text{ para } A \text{ implica } B,$$
$$A \leftarrow B \text{ por } A \text{ está implicada en } B,$$
$$A \mid B \text{ por } A \text{ incompatible con } B.$$

Los símbolos introducidos son utilizados por algunos autores, que escriben sobre lógica simbólica *. En esta anotación, las dos fórmulas

$$A \rightarrow B, \quad B \leftarrow A$$

son exactamente equivalentes, y lo mismo las fórmulas

$$A \mid B, \quad B \mid A.$$

También abreviaremos «digna de crédito» por «cr.» y «algo» por «a». Ver tabla I.

Tabla I

| | (1) Demostrativo | (2) Matizado demostrativo | (3) Matizado inductivo | (4) Inductivo |
|---|---|---|---|---|
| 1. Examinando una consecuencia | $A \rightarrow B$<br>$B$ falsa | $A \rightarrow B$<br>$B$ menos cr. | $A \rightarrow B$<br>$B$ más cr. | $A \rightarrow B$<br>$B$ verdadera |
| | $A$ falsa | $A$ menos cr. | $A$ a. más cr. | $A$ más cr. |
| 2. Examinando un fundamento posible | $A \leftarrow B$<br>$B$ verdadera | $A \leftarrow B$<br>$B$ más cr. | $A \leftarrow B$<br>$B$ menos cr. | $A \leftarrow B$<br>$B$ falsa |
| | $A$ verdadera | $A$ más cr. | $A$ a. menos cr. | $A$ menos cr. |
| 3. Examinando una conjetura conflictiva | $A \mid B$<br>$B$ verdadera | $A \mid B$<br>$B$ más cr. | $A \mid B$<br>$B$ menos cr. | $A \mid B$<br>$B$ falsa |
| | $A$ falsa | $A$ menos cr. | $A$ a. más cr. | $A$ más cr. |

* D. Hilbert y W. Ackerman, *Grundzüge der theoretischen Logik.*

## 8. Combinación de patrones sencillos.

En la investigación matemática puede surgir fácilmente la siguiente situación. Investigamos un teorema $A$. Este teorema $A$ ha sido formulado claramente, pero no sabemos si es verdadero o falso y deseamos saberlo. Después de algunos momentos damos con un posible fundamento: vemos que $A$ puede derivarse de otro teorema $H$.

$A$ está implicada en $H$

y así intentamos probar $H$. No lo logramos, pero observamos que una de sus consecuencias, $B$, es verdadera. La situación es:

$A$ implicada en $H$
$B$ implicada en $H$
$B$ verdadera

¿Se obtiene alguna inferencia concerniente a $A$ a partir de estas premisas?

Creo que hay una, que podemos obtener combinando dos de los patrones vistos en la sección anterior. En efecto, el patrón fundamental inductivo da por resultado:

$H$ implica $B$
$B$ verdadera
$H$ más digna de crédito.

Obteniendo esta conclusión hemos utilizado sólo dos de nuestras tres premisas. Vamos a combinar la tercer premisa que no hemos utilizado con la conclusión que obtuvimos; uno de nuestros patrones matizados (en la intersección de la segunda fila con la segunda columna de la tabla I en la sección 7) da:

$A$ implicada en $H$
$H$ más digna de crédito
A más digna de crédito.

El resultado es (como debía serlo) bastante obvio en sí mismo: a partir de la verificación de la consecuencia $B$ de su posible fundamento $H$ algún crédito se refleja sobre la misma proposición $A$,

## 9. Sobre la inferencia por analogía

La situación discutida en la sección anterior puede ser construida como un lazo entre los patrones discutidos en este capítulo y una de las más notables formas de inferencia plausible: la conclusión por analogía.

No creo que sea posible explicar la idea de analogía en términos completamente definidos de lógica formal; en cualquier caso, no tengo la intención de hacerlo. Como hemos visto antes, en la sección 2.4, la analogía tiene que ver con la semejanza y con las intenciones del pensador. Si notamos alguna semejanza entre dos objetos (o, preferiblemente, entre dos sistemas de objetos) e intentamos reducir esta semejanza a conceptos definidos, estamos pensando analógicamente.

Por ejemplo, observamos alguna semejanza entre dos teoremas $A$ y $B$; observamos algunos puntos comunes. Quizá pensamos que algún día será posible imaginar un teorema $H$ más comprensivo que reúna todos los puntos comunes esenciales y del cual se sigan naturalmente $A$ y $B$. Si empezamos a pensar así estamos pensando analógicamente *.

De cualquier modo, consideremos la analogía de dos teoremas $A$ y $B$ como la intención de descubrir un fundamento común $H$ a partir del cual pueden deducirse $A$ y $B$:

$A$ implicada en $H$, $\qquad$ $B$ implicade en $H$.

No olvidemos que no *tenemos* $H$; solamente *esperamos* que exista una tal $H$.

De pronto logramos probar uno de los dos teoremas análogos, digamos $B$. ¿Cómo influye este suceso sobre nuestra confianza en el otro teorema $A$? La situación tiene algo en común con la situación considerada en la anterior sección 8. Alcanzamos allí una razonable conclusión expresada por el patrón compuesto

$A$ implicada en $H$
$B$ implicada en $H$
$B$ verdadera
---
$A$ más digna de crédito.

* Así, el teorema isoperimétrico y la conjetura de Rayleigh, comparados entre sí en la sección 12.4, puede sugerirnos la idea de una generalización común.

Hay, por supuesto, una inferencia importante: que ahora no tenemos $H$, sólo la esperamos. Con esta precaución podemos mirar las dos premisas

$A$ implicada en $H$

$B$ implicadea en $H$

como equivalente a:

$A$ análoga a $B$.

Al sustituir esta premisa por las otras dos del anterior patrón compuesto llegamos al patrón fundamental de inferencia plausible primeramente exhibido en la sección 12.4:

$$\frac{\begin{array}{c}A \text{ análoga a } B \\ B \text{ verdadera}\end{array}}{A \text{ más digna de crédito.}}$$

## 10. Inferencia cualificada

Regresemos otra vez al patrón fundamental inductivo. Es el primer patrón que hemos presentado y constituye la forma más notable de razonamiento plausible. Se refiere a la verificación de la consecuencia de una conjetura y al cambio resultante en nuestra opinión. Nos dice algo siempre sobre la dirección de este cambio; una verificación tal sólo puede aumentar nuestra confianza en la conjetura. Pero no nos dice nada sobre la cantidad de cambio; el aumento de confianza puede ser grande o pequeño. Incluso puede llegar a ser tremendamente grande o ridículamente pequeño.

El principal objeto de la presente sección es aclarar las circunstancias de que dependen diferencias tan importantes. Empecemos por recordar uno de nuestros ejemplos (sección 12.3).

Un demandado es acusado de haber volado el yate del padre de su novia y la acusación presenta un recibo firmado por el demandado reconociendo la compra de tal y tal cantidad de dinamita.

La evidencia contra el demandado parece ser muy fuerte. ¿Por qué tanto? Insistamos sobre los rasgos generales del caso. Dos enunciados juegan un papel esencial,

$A$: el demandado voló el yate.

$B$: el demandado adquirió explosivos.

Al principio del proceso el tribunal tiende a considerar $A$ como una conjetura. La acusación trata de hacer $A$ más digno de crédito al jurado, la defensa trata de hacerlo menos digno de crédito.

Al principio del proceso también $B$ tiende a ser considerado como una conjetura. Más tarde, tras la presentación de ese recibo al tribunal (la autenticidad de la firma no fue rehusada por la defensa), $B$ tiende a ser considerado como un hecho probado.

Ciertas relaciones entre $A$ y $B$, no obstante, parecerán claras desde el principio.

$A$ sin $B$ es imposible. Si el demandado voló el yate tenía explosivos. De algún modo él habría adquirido estos explosivos: por compra, robo, regalo, herencia o cualquier otra manera. Es decir,

$A$ implica $B$.

$B$ sin $A$ no es imposible, pero parece extremadamente inverosímil desde el principio. Comprar dinamita no es muy corriente en un ciudadano normal. Comprar dinamita sin la intención de atentar contra algo o alguien parece absurdo. Es fácil imaginar que el demandante tuviese bastante fundamento emocional y económico para volar dicho yate. Es difícil sospechar cualquier otro propósito para la compra de dinamita, excepto volar el yate. Y así, como hemos dicho, $B$ sin $A$ parece ser extremadamente inverosímil.

Retengamos este importante componente de la situación: *la credibilidad de $B$, antes del suceso, disminuye bajo la asunción de que $A$ no es cierto.* Abreviemos esta precisa, aunque larga, descripción como «la credibilidad de $B$ sin $A$». Así podemos decir:

$B$ sin $A$ es poco digna de crédito.

Ahora podemos ver las premisas esenciales y el patrón total de la inferencia plausible que nos impresionó con su:

$A$ implica $B$

$B$ sin $A$ es poco digna de crédito

$B$ es verdadera

---

$A$ es mucho más digna de crédito.

Para mejor comprensión imaginemos que ese importante componente de la situación, la credibilidad de $B$ sin $A$, cambia gradualmente, varía entre sus casos extremos de modo continuo.

$A$ implica $B$. Si, inversamente, también $B$ implica $A$, de modo que

*A* y *B* se implican mutuamente, la credibilidad de *B* sin *A* alcanza su mínimo, es nula. En este caso, si *B* es cierta, *A* también es cierta, de modo que la credibilidad de *A* alcanza su máximo.

*A* implica *B*. Es decir, *B* es cierta cuando *A* es cierta. Si la credibilidad de *B* sin *A* se aproxima a su máximo, *B* es casi cierta cuando *A* es falsa. Por tanto, *B* es casi cierta en cualquier caso. Cuando ocurre un suceso que parece casi cierto por adelantado, no obtenemos mucha información nueva y así no podemos sacar consecuencias sorprendentes. (La compra de una hogaza de pan, por ejemplo, apenas puede rendir una tan fuerte evidencia circunstancial como la compra de dinamita.)

Asumamos, como en la sección 6, que la fuerza de la conclusión varía *continuamente en la misma dirección* cuando el factor influyente, la credibilidad de *B* sin *A*, cambia sin cambio de dirección. Entonces estos dos deben variar en direcciones opuestas y llegamos así a una importante calificación de la fortaleza de la conclusión en el patrón fundamental inductivo:

*A* implica *B*

*B* es verdadera

---

*A* es más digna de crédito.

*La fuerza de la conclusión aumenta cuando la credibilidad de B sin A disminuye.*

Pongamos uno al lado del otro los dos casos extremos:

| | |
|---|---|
| *A* implica *B*<br>*B* sin *A* es poco digna de crédito<br>*B* es verdadera<br>---<br>*A* es mucho más digna de crédito. | *A* implica *B*<br>*B* es casi siempre verdadera<br>*B* es verdadera<br>---<br>*A* es muy poco más digna de crédito. |

Las primeras dos premisas están entre paréntesis para expresar que la segunda es considerada como una cualificación de la primera. Que *A* implica *B* es la primera premisa en el patrón fundamental inductivo. Cualificamos esta premisa; añadimos una modificación que ejerce mucho peso en la determinación de la fortaleza de la conclusión. Por el gusto de comparar recordemos que en la sección 6 hemos modificado el patrón fundamental inductivo en otra dirección, debilitando su segunda premisa.

## 11. Sobre verificaciones sucesivas

Hemos verificado ya $n$ consecuencias $B_1, B_2, \ldots, B_n$ de una cierta conjetura $A$. Ahora probamos una nueva consecuencia $B_{n+1}$, ensayamos y encontramos que también $B_{n+1}$ es verdadera. ¿Cuál es la influencia de esta nueva evidencia sobre nuestra confianza en $A$? Por supuesto,

$$\frac{\begin{array}{c}A \text{ implica } B_{n+1} \\ B_{n+1} \text{ es verdadera}\end{array}}{A \text{ es más digna de crédito.}}$$

Sin embargo, ¿qué fuerza tiene esta conclusión? Eso depende de la credibilidad de $B_{n+1}$ sin $A$ como hemos visto en la sección anterior.

Ahora podíamos haber tenido una buena razón para creer en $B_{n+1}$ antes de que fuese verificado, aún bajo la asunción de que $A$ no es cierto. Hemos visto previamente que $B_1, B_2, \ldots, B_n$ son verdaderos. Si $B_{n+1}$ es muy semejante a $B_1, B_2, \ldots, B_n$ podemos predecir *por analogía* que también $B_{n+1}$ será cierto. Si $B_{n+1}$ es muy diferente de $B_1, B_2, \ldots, B_n$, no resulta apoyada por la analogía y tenemos muy poca razón para creer en $B_{n+1}$ sin $A$. Por tanto, *la fuerza de la confianza añadida resultante de una añadida verificación aumenta cuando la analogía de la consecuencia nuevamente verificada disminuye con respecto a las consecuencias previamente verificadas.*

Esto expresa esencialmente lo mismo que los patrones complementarios formulados en la sección 12.2, aunque quizá un poco mejor. En efecto, podemos considerar la mención explícita de la analogía como una ventaja.

## 12. Sobre conjeturas rivales

Si hay dos conjeturas diferentes, $A$ y $B$, presentadas para explicar el mismo fenómeno, las consideramos como opuestas la una de la otra, aunque no se haya probado que sean lógicamente incompatibles. Estas conjeturas $A$ y $B$ pueden o no pueden ser incompatibles, pero una de ellas tiende a hacer a la otra superflua. Esta es una oposición suficiente, y consideramos, por ello, $A$ y $B$ como conjeturas *rivales*.

Hay veces que tratamos con conjeturas rivales *como si* fuesen incompatibles. Por ejemplo, tenemos dos conjeturas rivales $A$ y $B$, pero a pesar de algún esfuerzo hecho no podemos pensar en una tercera

conjetura que explique el mismo fenómeno; entonces cada una de las dos conjeturas *A* y *B* es el «único rival obvio» de la otra. Una corta ilustración esquemática puede aclararnos el significado del término.

Supongamos que *A* es la teoría corpuscular de la luz que se remonta a Newton y que *B* es la teoría ondulatoria de la luz que se debe a Huyghens. Imaginemos también que discutimos estas cuestiones en tiempo posterior a Newton y Huyghens, aunque anterior a Youngs y Fresnel, cuando muchas discusiones no convincentes de estas teorías tuvieron lugar. Nadie mostró, o ha pretendido mostrar, que estas dos teorías son lógicamente incompatibles, y menos aún que sean las únicas alternativas lógicamente posibles; pero nunca hubo otras teorías prominentes de la la luz, a pesar de que los físicos tuvieron amplia oportunidad de inventar tales teorías: cada una de ellas fue la única rival obvia de la otra. Y así, toda teoría que pareciese atacar una de las dos rivales era interpretada como una defensa de la otra.

En general, la relación entre conjeturas rivales es semejante a la relación entre rivales en cualquier otra clase de competición. Si luchamos por un premio, la debilidad de cualquier rival nuestro significa algún fortalecimiento de nuestra posición. No ganamos mucho por el ligero retraso de uno de nuestros muchos y oscuros rivales. Ganamos algo más si tal retraso le ocurre a un peligroso rival. Ganamos todavía más si nuestro más peligroso rival se retira o queda fuera de la carrera. Si tenemos un rival único cualquier debilitación o fortalecimiento de su posición influye apreciablemente en nuestra posición. Y algo semejante ocurre entre conjeturas en competencia. Hay un patrón de razonamiento plausible que intentamos hacer algo más explícito en la tabla II.

TABLA II

| | |
|---|---|
| *A* incompatible con *B*<br>*B* falsa<br>———<br>*A* más digna de crédito | *A* incompatible con *B*<br>*B* menos digna de crédito<br>———<br>*A* algo más digna de crédito |
| *A* rival de *B*<br>*B* falsa<br>———<br>*A* un poco más digna de crédito. | *A* rival de *B*<br>*B* menos digna de crédito<br>———<br>*A* muy poco más digna de crédito. |

La disposición de la tabla II se explica casi por sí misma. Esta tabla consta de cuatro patrones ordenados en dos filas y dos columnas. La primera fila contiene dos patrones ya considerados; ver la sección 3, el fin de la sección 6 y la última fila de la tabla I. Al pasar de la primera fila a la segunda debilitamos la primera premisa; de hecho, sustituimos

una clara relación de lógica formal entre *A* y *B* por una relación algo difusa, aunque con un cierto sentido práctico. Esta debilitación de la primera premisa hace a la conclusión más débil en la misma medida, como la expresión verbal deja traslucir. Al pasar de la primera columna a la segunda debilitamos la segunda premisa, que hace a la conclusión más débil en la misma medida. El patrón de la esquina sudeste no tiene ninguna premisa con carácter lógicamente demostrativo y por ello su conclusión es la más débil.

Es importante recalcar que las expresiones verbales utilizadas son ligeramente desorientadoras. En efecto, las especificaciones añadidas a «más digno de crédito» («algo», «un poco», «muy poco») no transmiten un *absoluto*, sino sólo un *relativo*, grado de credibilidad. Indican solamente el cambio de fuerza lógica cuando pasamos de una fila a otra, de una a otra columna. Aun así, el patrón más débil de los cuatro puede darnos una conclusión de peso si la convicción de que la conjetura *A* no tiene rival más peligroso que *B* es lo bastante fuerte. Este papel jugará algún papel en el próximo capítulo.

## 13. Sobre la prueba judicial

El razonamiento por el que un tribunal llega a sus decisiones puede ser comparado con el razonamiento inductivo en el que el naturalista apoya sus conclusiones. Tales comparaciones han sido ya hechas y debatidas por autoridades sobre procedimiento legal *. Empecemos a discutir este interesante punto considerando un ejemplo.

(1) El gerente de un popular restaurante que permanece abierto hasta altas horas volvió a su casa particular, como era usual en él, bien entrada la medianoche. Cuando salió de su coche para abrir la puerta del garaje fue atracado y robado por dos individuos enmascarados. La policía, buscando las premisas, encontró un oscuro trapo gris en el patio delantero de la víctima; el trapo debió ser usado por uno de los atracadores para cubrir su rostro. La policía interrogó a varias personas de la vecindad. Uno de los hombres interrogados tenía un abrigo con un agujero muy grande, aunque por otro lado el paño estaba en buenas condiciones. El trapo encontrado en el patio anterior de la víctima del atraco era del mismo material que el paño del abrigo y se adecuaba perfectamente a la forma del agujero. El hombre del abrigo fue arrestado e inculpado de participación en el atraco.

---

* J. H. Wigmore, *The principles of judicial proof*, Boston, 1913; cf. págs. 9-12, 15-17.

(2) Muchos de nosotros estaremos concordes en que una inculpación semejante estaba ampliamente justificada por las circunstancias referidas. Pero, ¿por qué? ¿Cuál es la idea subyacente?

La acusación no es un enunciado de hechos, sino la expresión de una sospecha, de una *conjetura:*

*A.* El hombre del abrigo participó en el atraco.

No obstante, una acusación oficial semejante no puede ser una conjetura gratuita, sino apoyada en hechos relevantes. La conjetura *A* es apoyada por el hecho.

*B.* El trapo encontrado en el patio anterior de la víctima del atraco es del mismo material que el paño del abrigo del acusado y, por si fuera poco, se adapta exactamente al agujero que existe en dicho abrigo.

Sin embargo, ¿por qué consideramos *A* como una justificación de *B*? No olvidemos que *A* es sólo una conjetura: puede ser verdadera o falsa. Si deseamos actuar honradamente, debemos considerar cuidadosamente ambas posibilidades.

Si *A* es cierto, *B* parece realmente incomprensible. Podemos fácilmente imaginar que un hombre con urgente necesidad de enmascararse y no teniendo a su alcance cosa mejor arranca un trozo del paño de su abrigo. Con las prisas para abandonar el lugar del delito este hombre pierde su máscara. O, con el susto, pudo arrojar su máscara allí en lugar de guardarla en el bolsillo o ponerla en otro lugar a salvo. En resumen, *B* con *A* parece más digno de crédito.

Si, no obstante, *A* no es verdadero, *B* parece inexplicable. Si el hombre no participó en un atraco o algo así, ¿por qué estropea su abrigo en perfectas condiciones, cortando una gran pieza del paño? Y ¿por qué, entre todos los lugares, encontramos dicha pieza sobre el mismo sitio que unos bandidos enmascarados han ejecutado su robo? Se podría tratar de una coincidencia, pero una coincidencia semejante resulta muy difícil de creer. En resumen, *B* sin *A* resulta muy poco digno de crédito.

Y así vemos que la conclusión que conduce a la acusación contra el hombre del abrigo tiene los siguientes patrones:

$$\left.\begin{array}{l} B \text{ con } A \text{ realmente digna de crédito} \\ B \text{ sin } A \text{ escasamente digna de crédito} \end{array}\right\}$$

$$\frac{B \text{ verdadera}}{A \text{ más digna de crédito.}}$$

Sin embargo, este patrón de razonamiento plausible está muy relacionado con otro patrón de razonamiento plausible que hemos discutido antes (en la sección 10):

$$\frac{\left.\begin{array}{l} A \text{ implica } B \\ B \text{ sin } A \text{ escasamente digna de crédito} \end{array}\right\} \qquad B \text{ verdadera}}{A \text{ mucho más digna de crédito.}}$$

La diferencia entre los dos patrones aparece desde el principio. La premisa

$$B \text{ con } A \text{ realmente digna de crédito}$$

es semejante a la premisa, aunque más débil que ella.

$$A \text{ implica } B,$$

que, de hecho, podría expresarse como «$B$ con $A$ cierto». Así, el primer patrón (descubierto) aparece como una «forma debilitada» del último patrón (presentado en la sección 10) y así eventualmente como una modificación del patrón fundamental inductivo (formulado en la sección 12.1).

El caso que nos ha conducido a la formulación del nuevo patrón ha sido muy sencillo. Veamos ahora un caso más complicado*.

(3) En el momento del asesinato, Clarence B. Hiller, con su esposa y cuatro niños, vivía en una casa de dos pisos de Chicago. Los dormitorios de la familia estaban en el segundo piso. Al principio de las escaleras que conducían al segundo piso había una luz de gas luciendo durante la noche. Un poco después de las dos de la madrugada del lunes la señora Hiller se despertó y notó que la luz no estaba. Inmediatamente despertó a su marido, que se dirigió a la escalera en batín de noche; allí encontró un intruso con el que luchó, cayendo ambos rodando a los pies de la escalera, donde Hiller fue tiroteado dos veces; murió a los pocos momentos. El tiroteo ocurrió sobre las 2,25 de la madrugada.

Justo un poco antes del tiroteo una de las hijas de Hiller había visto un hombre en la puerta de su dormitorio, sosteniendo una cerilla de modo que su cara no podía ser vista. Ella no se asustó, pues su padre acostumbraba a levantarse por la noche para ver si los niños estaban bien. Nadie en la familia vio al intruso.

---

* Sobre el caso siguiente el lector debe consultar los fallos de la Corte Suprema de Illinois, casi completamente reimpresos en Wigmore, l. c., nota 5, págs. 83-88.

A unos tres cuartos de milla de la casa de Hiller había una parada de autobuses. Muy temprano, la misma mañana que ocurrió el asesinato, cuatro policías que acababan de salir de servicio en la vecindad estaban sentados sobre un banco de esta parada, esperando el autobús. Sobre las 2,38 de la madrugada vieron un hombre que venía de una dirección desde la que el banco no podía ser visto fácilmente. Los oficiales hablaron al hombre, pero éste continuó con su mano derecha metida en el bolsillo; en vista de lo cual, pararon al hombre y le preguntaron. En su bolsillo llevaba un revólver cargado; estaba sudando; había sangre fresca en diferentes lugares de sus ropas; tenía una ligera herida en su antebrazo izquierdo, que sangraba un poco. Los oficiales (que no sabían nada en ese momento del asesinato) llevaron al hombre a la estación de policía, donde fue examinado. Este hombre —a quien llamarenos desde ahora el demandado— fue más tarde acusado del asesinato del señor Hiller.

El tribunal debía examinar el caso, y, tras el examen, negar o mantener la acusación, es decir, la acusación adelantó la siguiente conjetura:

*A*. El demandado disparó sobre el señor Hiller y le mató.

Examinemos los principales puntos de la evidencia adelantada en apoyo de la conjetura *A*.

$B_1$. Había pólvora quemada en dos recámaras del cilindro del revólver encontrado sobre el demandado en el momento de ser arrestado, y dicho revólver tenía también olor a humo fresco. A juicio de los oficiales de policía había sido usado dos veces en la hora anterior al arresto. Los cinco cartuchos con que el revólver estaba cargado ostentaban exactamente la marca de la misma fábrica que los tres cartuchos descargados encontrados en el pasillo de la casa de Hiller cerca del cuerpo del muerto.

$B_2$. El intruso entró en la casa de Hiller por una de las ventanas de guillotina de la cocina, después de quitar la tela metálica. Una persona que entrase por esta ventana podía apoyarse sobre la barandilla del porche. Sobre esta barandilla, que estaba recién pintada, había la huella de cuatro dedos de una mano izquierda. Dos empleados de la oficina de identificación de la policía de Chicago atestiguaron que, a su juicio, las huellas sobre la barandilla eran idénticas a las impresiones digitales del demandado.

$B_3$. Dos expertos que no pertenecían a la policía de Chicago expresaron la misma opinión sobre las huellas. (Uno era un inspector de la policía del Gobierno de Ottawa, Canadá; el otro, un primer experto del Gobierno federal de Washington, D. F.)

$B_4$. Sobre las dos de la madrugada, justo un poco antes del tiroteo

al señor Hiller, un vagabundo entró en una casa separada de la de Hiller por un solar vacío. Dos mujeres vieron a un hombre en la puerta de su dormitorio con una cerilla encendida sobre su cabeza. Ambas mujeres atestiguaron que este vagabundo era del mismo tamaño y forma que el demandado. Una de las mujeres recordaba que el hombre vestía una camisa de color rojo y unos calzones floreados. Habiendo inspeccionado la camisa y los calzones del demandado presentados en el tribunal, la testigo atestiguó que en su opinión el demandado era el mismo hombre que vio por la noche en la puerta.

$B_5$. El demandado dio un nombre falso y una falsa dirección cuando fue arrestado y negó que hubiese sido arrestado antes. De hecho, él había sido detenido antes bajo la acusación de robo con escalamiento y libertado bajo palabra; más tarde volvió a la penitenciaría por violación de la palabra y soltado una segunda vez bajo palabra seis semanas antes de la noche del asesinato. Dos semanas después de su segunda palabra compró un revólver con nombre falso, lo empeñó, lo desempeñó, lo volvió a empeñar y lo desempeñó por segunda vez cinco horas antes del asesinato.

$B_6$ El demandado fue incapaz de explicar razonablemente la sangre de sus ropas, la herida de su antebrazo izquierdo y sus andanzas en la noche del tiroteo. Sobre sus andanzas él hizo dos relatos diferentes, una después de su arresto y otra en el tribunal. Las gentes a quienes dijo haber visitado esa noche negaron haberle visto. Entonces el demandado dijo al tribunal que había estado en un bar, pero no pudo encontrarse ningún testigo que lo corroborara.

(4) Todos los hechos, sucesos y circunstancias relatados bajo los encabezamientos $B_1$, $B_2$, ... $B_6$ son realmente comprensibles *si* la acusación $A$ es cierta. Todos apoyan $A$, aunque el peso de tal apoyo no es el mismo en todos los casos. Algunos de estos hechos serían igualmente explicables aun cuando $A$ no fuese cierto. No obstante, otros aparecen como milagrosas coincidencias si $A$ no fuese cierto.

Que los cartuchos encontrados en el revólver del demandado sean de la misma marca que los encontrados junto al cuerpo de la víctima prueba poco en sí mismo, si esta marca es usualmente vendida por todos los armeros. Sin embargo, prueba mucho que el número de disparos hechos con este revólver coincida con los tiros recibidos por la víctima y dentro de la misma hora; es muy difícil explicar una coincidencia semejante. La coincidencia entre las huellas sobre la barandilla y las huellas del demandado sería considerada como una prueba decisiva hoy en día, pero no ocurría lo mismo en la época del juicio, el año 1911. Que el demandado falsee su nombre, su dirección y su pasado criminal no prueba

mucho; tales mentiras son comprensibles si la acusación $A$ es cierta, pero también si no lo es: el hombre quiere que la policía le deje tranquilo en cualquier caso. Sin embargo, es de mucho peso que el demandado no pueda dar una explicación consistente de sus andanzas en la noche fatídica. El debía conocer la importancia de este punto, y su asesor sabía sin duda la importancia de una coartada. Si la acusación $A$ fuese falsa y el demandado hubiese pasado la noche sin pena ni gloria o cometiendo un delito menor, ¿por qué no decirlo todo, al menos antes de que fuese demasiado tarde?

Todos los detalles mencionados son realmente comprensibles si la acusación $A$ es cierta. Pues la coincidencia de tantos detalles resulta inexplicable si la acusación $A$ no es cierta; es totalmente increíble que tantas coincidencias se deban sólo al azar. En cualquier caso, la defensa fracasó en presentar una explicación convincente de la evidencia sometida.

Había, por supuesto, más evidencias de las presentadas aquí, y había una cuya descripción no puede hacerse adecuadamente: la conducta del demandado y de los testigos. El jurado estaba convencido de que el demandado había cometido el crimen y el Tribunal Supremo mantuvo la convicción. A continuación citamos la última sentencia del presidente de la Sala: «Ninguna de estas circunstancias, consideradas por separado, sería concluyente de su culpabilidad [la del demandado], pero cuando todos los hechos y circunstancias presentadas en evidencia se dan juntas, el jurado se siente justificado a creer que el veredicto de culpabilidad sigue una consecuencia lógica» *.

(5) Los enunciados $B_1$, $B_2$, ..., $B_6$ enumerados bajo (3) se adecuan bastante bien al patrón presentado en (2). Pero se adecuan aún mejor a otro patrón que difiere de aquél en un solo punto:

$$\left.\begin{array}{l} B \text{ con } A \text{ realmente digna de crédito} \\ B \text{ sin } A \text{ realmente menos digna de crédito} \end{array}\right\}$$
$$\frac{B \text{ verdadera}}{A \text{ más digna de crédito.}}$$

Cada uno de los enunciados $B_1$, $B_2$, ..., $B_6$ puede ser significativamente sustituido por $B$ en este patrón en que $A$ será interpretado como la acusación. Los enunciados $B_1$, $B_2$, ..., $B_6$ son enunciados compuestos; tienen partes (algunas de las cuales hemos enfatizado en (4)) cada una de las cuales puede ser considerada en sí misma como evi-

* Cf. WIGMORE, *loc. cit.*, 5, pág. 88.

dencia relevante: cada porción puede ser significativamente sustituida en el patrón de arriba. Si recordamos nuestra discusión bajo (4), podemos comprobar que las $B$ menos fácilmente creíbles no necesitan $A$, la conclusión es la más fuerte.

Si viéramos cómo los testigos sucesivos desplegaban la masa de evidencias ante el jurado en el curso del procedimiento, podríamos comprender más claramente los papeles análogos del razonamiento plausible en tales procedimientos y en una investigación científica en que varias consecuencias de una conjetura son sucesivamente intentadas. (Cf., especialmente, la sección 12.2.)

(6) La consideración anterior claramente sugiere un patrón compuesto de razonamiento plausible que está relacionado con el patrón establecido en (5) justo como el patrón compuesto presentado en la sección 12.2 está relacionado con el patrón fundamental inductivo de la sección 12.1. No entramos aquí en esta cuestión; un lector más versado en la doctrina de la evidencia legal podría darnos ejemplos e interpretaciones más impresionantes, pero por nuestra parte no añadiremos ninguna otra ilustración del patrón *.

Cuando Colón y sus compañeros navegaban hacia el Occidente a través del Océano desconocido, se sintieron animados al ver pájaros. Consideraron los pájaros como un signo favorable, que indicaba la proximidad de tierra. Aunque fueron repetidamente desilusionados, el razonamiento subyacente me parece bastante correcto. Formulado en longitud este razonamiento dice así:

Cuando el barco está cerca de tierra, vemos pájaros
Cuando el barco está lejos de tierra, vemos pájaros menos a menudo

Ahora vemos pájaros

Por tanto, se hace más digno de crédito que estamos cerca de tierra.

Este razonamiento se adecua perfectamente al patrón formulado en (5): la presencia de pájaros es considerada como circunstancial evidencia de la cercanía de tierra. La tripulación de Colón vio varios pájaros el jueves, 11 de octubre de 1492, y al día siguiente descubrieron la primera isla de un Nuevo Mundo.

El lector puede observar que el patrón ilustrado subyace en muchos de nuestros razonamientos cotidianos.

---

* *How to Solve It*, págs. 212-221.

## EJEMPLOS Y COMENTARIOS

### Primera parte

**1.** Siguiendo el método de la sección 4 (6), derivar el patrón demostrativo mencionado en la sección 3 del patrón demostrativo mencionado en la sección 1.

**2.** Dar los detalles de la demostración esbozada en la sección 5: derivar el patrón heurístico de la sección 2 a partir de la sección 1.

**3.** Derivar el patrón heurístico de la sección 3 del de la sección 1.

**4.** En un crucigrama encontramos una palabra de nueve letras, y la clave es: «Disagreeable form of tiredness» *.

La condición que la palabra desconocida tiene que satisfacer está formulada muy ambiguamente, por supuesto. Después de algunos intentos sin éxito, observamos que «tiredness» tiene nueve letras, precisamente tantas como la palabra desconocida, y esto puede conducirnos a la siguiente conjetura:

*A*. La palabra desconocida significa «disagreeable» y es un *anagrama* de TIREDNESS.

(Anagrama significa la reordenación de las letras de la palabra dada dentro de un nuevo orden.) Esta conjetura *A* puede parecer bastante verosímil. (En efecto, «forma de» puede sugerir «anagrama de» en la jerga de los crucigramas.) Tomando otras palabras desconocidas del rompecabezas encontramos soluciones bastante plausibles para dos de las palabras que cruzan las nueve letras de la palabra mencionada, que nos dan dos posibles letras colocadas como indica el siguiente diagrama:

$$— — — — — — T — R.$$

Podemos considerar esto como una evidencia de nuestra conjetura *A*.

(*a*) ¿Por qué? Saquemos el patrón apropiado.

(*b*) Intentar encontrar las nueve letras de la palabra buscada. [Al hacerlo así tenemos una oportunidad de pesar la evidencia de *A*. He aquí algunas preguntas que nos ayudarán: ¿Qué letra es la más verosímil

---

* *The Manchester Guardian Weekle*, 29 de noviembre de 1951. (Este juego de palabras es imposible de traducir al español; de aquí que conservemos las palabras claves en inglés. Para ayuda del lector, tenga en cuenta que «disagreeable» significa desagradable y «tiredness» significa cansancio. También hemos conservado la palabra inglesa en la solución a este problema al final del libro.—*N. del T.)*

entre *T* y *R*? ¿Cómo pueden ser colocadas las vocales *E E I*? Ver tambien *How to Solve It*, págs. 147-149.]

**5.** Tomemos de nuevo el caso judicial ya considerado en la sección 12.3 y (más concretamente) en la sección 10. Consideremos de nuevo la acusación (el *factum probandum*, el hecho a ser probado de la acusación):

*A*. El demandado voló el yate.

No obstante, cambiemos nuestras notas en otro punto: consideremos aquí el enunciado

*B*. El demandado compró dinamita en tal y tal cantidad, en tal y tal tienda, tal y tal día.

El cambio es que *B* denota aquí no un enunciado general, sino un hecho específico. (Los tribunales prefieren, o deben preferir, tratar con hechos tan distintamente específicos como sea posible.) De nuevo tomamos *B* como si estuviera probado. (Así, *B* es un *factum probans*, un hecho que apoya la prueba.)

Tal cambio de notación no puede cambiar la fuerza del argumento. Sin embargo, ¿cuál es ahora el patrón?

**6.** Los demandados son ahora un contratista y un oficial público. El primero es acusado de sobornar al segundo y éste de aceptar el soborno. Hay un cargo específico: la acusación afirma que el pago al contado del nuevo coche del oficial salió del bolsillo del contratista. Uno de los testigos de la prosecución es un vendedor de coches, el cual testificó que el 29 de noviembre recibió 875 dólares como pago por el coche del oficial. Otro testigo es el gerente de un Banco local, quien atestiguó que el 27 de noviembre (del mismo año) fueron retirados 875 dólares de una cuenta inactiva conjunta del contratista y su esposa; el recibo estaba firmado por la esposa. Estos hechos no fueron recusados por la defensa.

¿Cuál considera usted que es el punto fuerte en esta evidencia? Dé el patrón apropiado.

**7.** Los Blacks, los Whites y los Greens viven en la misma calle de Suburbia. Los Blacks y los Whites son vecinos y los Greens viven precisamente en el lado opuesto. Una tarde el señor Black y la señora White tenían una larga conversación sobre la valla de su jardín. Era bastante oscuro, pero la señora Green no dejó de observar la conversación y sacó la conclusión; ya sabemos la conclusión: la conjetura preferida por la señora Green.

Desgraciadamente, no hay mucha oportunidad de oponerse a la charlatanería de la señora Green. Si, no obstante, desafiando un peligro

terrible, deseamos convertirnos en defensores de las personas inocentes calumniadas por la señora Green, yo puedo darle un hecho: los Whites, que desean mudarse cerca de la oficina del padre de familia por una larga temporada, firmaron un arrendamiento de una casa que pertenecía al tío del señor Black, y esto sucedió unos pocos días después de la citada conversación. Utilicemos este hecho.

¿Cuál es su defensa y cuál es el patrón?

**8.** La acusación intenta probar:

*A.* El demandado conocía, y fue capaz de reconocer, a la víctima en el momento del crimen.

La acusación apoya esto por hecho innegable:

*C.* El demandado y la víctima estuvieron ambos durante varios meses empleados por la misma firma tres años antes de que el crimen fuese cometido.

Así, *A* es el *factum probandum* y *C* es adelantado como *factum probans*. ¿Cuál es el patrón? [La notación está ingeniada para ayudarle. ¿Es relevante el tamaño de la firma?]

**9.** *Sobre la investigación inductiva en matemáticas y en las ciencias físicas.* Una diferencia entre «situaciones matemáticas» y «situaciones físicas» que parece ser importante desde el punto de vista del razonamiento plausible ha sido apuntado en la sección 6. Hay también otras diferencias de esta clase, y una de ellas será discutida aquí.

Coulomb descubrió que la fuerza entre cuerpos electrificados varía inversamente al cuadrado de la distancia. El apoyó esta ley del cuadrado inverso en experimentos directos con la balanza de torsión. Los descubrimientos de Coulomb eran delicados y la discrepancia entre sus cantidades teóricas y experimentales, considerable. Nosotros no podemos dejar de sospechar que sin la poderosa analogía de la ley de Newton (la ley del cuadrado inverso en las atracciones gravitatorias), ni Coulomb mismo ni sus contemporáneos habrían considerado sus experimentos con la balanza de torsión como concluyentes.

Cavendish descubrió la ley del cuadrado inverso de las atracciones y repulsiones eléctricas independientemente de Coulomb. (Las investigaciones de Cavendish no fueron publicadas durante su vida, y la prioridad de Coulomb resulta incontestable.) Sin embargo, Cavendish ingenió un experimento más sutil en apoyo de la ley. No necesitamos discutir los detalles de su método *, sólo uno de sus rasgos es esencial aquí: Cavendish consideró la posibilidad de que la intensidad de la fuerza no fuese

---

* Cf. J. C. MAXWELL, *A treatise on electricity and magnetism*, 2.ª ed. (1881), volumen 1, págs. 76-82.

$r^{-2}$ ($r$ es la distancia de las cargas eléctricas), sino, más generalmente, $r^{-\alpha}$ donde $\alpha$ es una constante positiva. Su experimento muestra que $\alpha - 2$ puede exceder en valor absoluto a una cierta fracción numérica.

La investigación experimental de Coulomb es bastante semejante a una investigación inductiva en matemáticas: él confronta consecuencias de una ley física conjetural con las observaciones del mismo modo que el matemático confronta consecuencias particulares de una conjetural ley teórico-numérica con las observaciones. La analogía juega un importante papel en la elección de la ley conjetural, tanto aquí como allí. Sin embargo, la investigación experimental de Cavendish es de un carácter diferente; no sólo considera éste una ley conjetural (la ley $r^{-2}$), sino varias ($r^{-\alpha}$). Estas leyes son diferentes (diferentes leyes de atracción eléctrica corresponden a diferentes valores del parámetro $\alpha$), pero están relacionadas, pues pertenecen a la misma «familia» de leyes. Cavendish confronta a la familia total de leyes con las observaciones e intenta extraer la ley que mejor concuerda con ellas.

Esta es la diferencia más característica entre dos investigaciones: una apunta a *una* conjetura, la otra a una *familia* de conjeturas. La primera compara las observaciones con las consecuencias de una conjetura, la otra las compara con las consecuencias de varias conjeturas simultáneamente. La primera intenta juzgar sobre las bases de tal comparación si la conjetura propuesta es aceptable o no, la otra intenta entresacar la conjetura más aceptable (o la menos inaceptable). La primera clase de investigación inductiva es muy practicada en matemáticas y no es corriente en ciencias físicas. La segunda clase de investigación matemática, por el contrario, es muy practicada en ciencias físicas, pero muy rara vez nos encontramos con ella en matemáticas.

En efecto, la clase más típica de experimento físico apunta a la medida de alguna constante física, para determinar su valor, como el experimento de Cavendish apunta a determinar el valor del exponente $\alpha$. En matemáticas podremos considerar esta o esa investigación como una investigación inductiva que apunta a la determinación de alguna constante matemática, pero tal tipo de investigaciones suelen ser bastante excepcionales *.

**10.** *Formulaciones generales tentativas.* El dicho a menudo repetido por Newton «Hipothesis non fingo» es unilateral. Sería un error interpretado como «Desconfiar de las conjeturas»: tal consejo, si fuese

---

* A propósito, el experimento de Cavendish tiene un horizonte más amplio: tiende a mostrar que la ley $r^{-2}$ es más aceptable que cualquier otra ley $\psi(r)$, sin restringir la función $\psi(r)$ a la forma $r^{-\alpha}$; cf. MAXWELL, *ibíd.*, págs. 76-82.

seguido, arruinaría la investigación inductiva. Es mejor consejo este otro: Ser rápido en formar conjeturas; lentos en aceptarlas. Todavía mejores son las palabras de Faraday: «El filósofo será un hombre que, escuchando todas las sugerencias, está determinado a juzgarlas por sí mismo». Por supuesto, el filósofo que Faraday tenía en la mente cultiva la filosofía experimental, no la tradicional.

El presente estudio intenta ser una investigación «inductiva» del razonamiento plausible. Yo presento aquí, sin inhibiciones, unas cuantas generalizaciones tentativas. Estas se aplicarán a unas pocas formas de razonamiento plausible, pero su formulación será cuidadosamente escudriñada antes de cualquier intento de extenderla más lejos.

(1) *Monotonicidad.* Las consideraciones de la sección 6 pueden sugerir una regla: «La conclusión de una inferencia plausible varía monotónicamente cuando una de sus premisas varía monotónicamente». Esto se adecua al caso considerado en la sección 6 y a unos pocos casos más, algunos de los cuales serán considerados ahora.

(2) *Continuidad.* Necesitaremos un término de lógica demostrativa. Decimos que

$A$ y $B$ son equivalentes

cuando $A$ y $B$ se implican una a otra mutuamente, es decir, cuando $A$ se sigue de $B$ y también $B$ se sigue de $A$. Si $A$ y $B$ son equivalentes, no podemos saber de momento si $A$ o $B$ es verdadero, pero sabemos que sólo dos casos son posibles: o ambos son verdaderos o ambos son falsos; $A$ y $B$ o caen o se mantienen juntos. Una expresión simbólica y descriptiva para la equivalencia de $A$ y $B$ es:

$$A \rightleftarrows B.$$

Las dos flechas sugieren que podemos pasar de la verdad de cualquiera de los dos enunciados $A$ y $B$ a la verdad del otro.

En la sección 10 consideramos una conexión sugestiva entre dos enunciados $A$ y $B$. Consideramos allí la relación lógica

$A$ implica $B$

lo mismo que la credibilidad de $B$ sin $A$. Imaginemos que esta credibilidad varía monotónicamente: $B$ se hace menos y menos digna de crédito sin $A$. En el límite cuando $B$ se hace imposible sin $A$, la verdad de $B$ implica la de $A$. Sin embargo, supusimos que la verdad de $A$ implica la de $B$, y así, en el límite, $A$ y $B$ implicadas mutuamente entre sí, se hacen equivalentes.

Observemos ahora cómo la variación descrita afecta a nuestro patrón fundamental inductivo:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ verdadera}\end{array}}{A \text{ más digna de crédito.}}$$

Asumimos que la segunda premisa permanece invariable cuando la primera premisa varía como hemos descrito. Como $B$ se hace menos y menos digno de crédito sin $A$, la conjetura $A$ se hace más y más digna de crédito por la verificación de su consecuencia $B$. Es decir, la conclusión se hace más fuerte, gana peso. En el límite la conclusión alcanza el «$A$ es verdadero», y así nuestro patrón de inferencia plausible alcanza en el límite el siguiente (obvio) patrón de inferencia demostrativa:

$$\frac{\begin{array}{c}A \text{ y } B \text{ equivalentes} \\ B \text{ verdadera}\end{array}}{A \text{ verdadera.}}$$

En resumen, nuestro patrón de inferencia plausible tiene una «forma límite», que es un patrón de inferencia demostrativa. Como las premisas de la inferencia plausible «tienden» a las premisas correspondientes de la forma límite, la conclusión plausible se «aproxima» en su extremo a la fortaleza límite. Aún más corto: hay una transición continua desde el patrón heurístico hasta el patrón demostrativo.

La mayor parte de esta descripción vale para unos cuantos casos más; algunos de ellos son desplegados en la tabla III. Los símbolos y abreviaturas de la tabla I explicados en la sección 7 son utilizados también en la tabla III. Igualmente, usamos el símbolo $\leftrightarrows$, explicado arriba. En lugar de no-$B$, definido en la sección 4, utilizamos el símbolo más corto $\bar{B}$.

TABLA III

| | $A \dashrightarrow B$ | $A \rightarrow B$ | $A \mid B$ | $A \mid B$ |
|---|---|---|---|---|
| | $B$ verdadera | $B$ más cr. | $B$ falsa | $B$ menos *cr.* |
| Aproximación..... | $A$ más cr. | $A$ a. más cr. | $A$ más cr. | $A$ a. más cr. |
| | $A \rightleftarrows B$ | $A \leftrightarrows B$ | $A \dashleftarrow \bar{B}$ | $A \rightleftarrows \bar{B}$ |
| | $B$ verdadera | $B$ más cr. | $B$ falsa | $B$ menos cr. |
| Límite..... | $A$ verdadera | $A$ más cr. | $A$ verdadera | $A$ más cr. |

(3) *¿A lo plausible desde lo demostrativo?* La tabla III puede traernos otra sugerencia. Los patrones límite de esta tabla son mucho más obvios que los de aproximación. Dos de los patrones límite son demostrativos, y los otros dos, aunque no puramente demostrativos, son poco problemáticos. Lo más debatible de los patrones aproximativos, que son todos los de inferencia plausible, parece provenir de los correspondientes patrones límite por un uniforme proceso de «debilitamiento»: los más fuertes enunciados como

$$A \rightleftarrows B, \quad A \text{ verdadera}, \quad A \text{ más cr.}$$

son sistemáticamente remplazados por los enunciados correspondientes más débiles

$$A \rightarrow B, \quad A \text{ más cr.}, \quad A \text{ a. más cr.}$$

¿Podrían *todas* las formas de inferencia plausible ser ligadas de algún modo análogo a las formas de inferencia demostrativa, o casi demostrativa?

## Segunda parte

El ejemplo 11 deberá ser leído en primer lugar, pues presenta (y justifica) lo que sigue.

**11.** *Más personal, más complejo.* En lo anterior no he discutido conjeturas de mi obra matemática publicada. Esto es una omisión, puesto que, después de todo, me resulta imposible conocer a ningún matemático más íntimamente que a mí mismo. Esta omisión les parecerá sospechosa a algunos lectores. No creo que tal sospecha esté justificada. La razón para no discutir temas más complejos y especializados, investigados por mí mismo, no es la falta de franqueza, sino precisamente la complejidad y especialización de los temas; me parece mejor discutir cuestiones más sencillas, de interés más general.

Los siguientes ejemplos, 12-19, suponen en el lector unos conocimientos más avanzados que los que requiere el presente libro. Están tomados de mis investigaciones. En ellos intento ofrecer una muestra representativa, en la que se incluyen algunas conjeturas que han aparecido ya impresas y otras aún no publicadas. Algunas de estas conjeturas pertenecen a mis primeros «ingenuos» trabajos, cuando aún no había empezado a pensar explícitamente en el tema del razonamiento plausible: otras pertenecen a mi obra posterior, menos ingenua. El ejemplo 12 procede de mis años ingenuos; habla de los fundamentos heurísticos que me llevaron a un resultado. Los ejemplos 13, 14, 15 y 16 tratan de conje-

turas primeramente publicadas por mí; el ejemplo 17, de una conjetura publicada al principio de mi obra menos ingenua, y los ejemplos 18 y 19 tratan de conjeturas que no han sido llevadas a la imprenta antes *.

Añadiré que, aun en mis años ingenuos, me encontraba impresionado y algo confuso por la fuerza de la confianza con que algunas de mis propias conjeturas me inspiraban, y me preguntaba qué clase de razones podía haber bajo tal confianza. Las siguientes frases expresan perfectamente mis primeras opiniones sobre el origen de nuevas conjeturas.

«Tenemos una línea recta que une dos puntos dados. Un nuevo teorema es a menudo una generalización que une dos casos particulares extremos, y se obtiene por una especie de interpolación lineal. Tenemos una línea recta con una dirección determinada a través de un punto dado. Un nuevo teorema es con frecuencia concebido en el afortunado momento en que una dirección general de investigación se encuentra con un caso particular apropiado. Un nuevo teorema puede también obtenerse desde el trazado de una paralela» **.

**12.** *Tenemos una línea recta que une dos puntos dados.* Los números $A_n^{(k)}$ dispuestos en el cuadrado infinito

$$\begin{array}{ccccccc} A_0^{(1)} & A_1^{(1)} & A_2^{(1)} & A_3^{(1)} & \dots & A_n^{(1)} & \dots \\ A_0^{(2)} & A_1^{(2)} & A_2^{(2)} & A_3^{(2)} & \dots & A_n^{(2)} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_0^{(k)} & A_1^{(k)} & A_2^{(k)} & A_3^{(k)} & \dots & A_n^{(k)} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array}$$

están en conexión con una función $f(x)$. (Son derivados del desarrollo de $f(x)$ en valores de $x$; tanto $A_n^{(k)}$ como $f(x)$ son reales.) E. Laguerre descubrió que $V(k)$, el número de cambios de signo en la fila $A_0^{(k)}$,

* En esta nota se citan sin nombre ensayos del autor de este libro; «cf.» indica la página en que se establece la conjetura (a veces en forma de pregunta); un ensayo citado con nombre llevó la primera prueba a la conjetura discutida. Ejemplo 12: *Rendiconti, Circolo Matematico di Palermo*, vol. 34 (1912), págs. 89-120. Ejemplo 13: *L'Intermédiaire des Mathématiciens*, vol. 21 (1914), pág. 27, qu. 4340; G. Szegö, *Math Annalen*, vol. 76 (1915), págs. 490-503. Ejemplo 14: *Math. Annalen*, vol. 77 (1916), páginas 497-513; cf. pág. 510; F. Carlson, *Math. Zeitschrift*, vol. 9 (1921), págs. 1-13. Ejemplo 15: *L'Intermédiaire des Mathématiciens*, vol. 20 (1913), págs. 145-146, qu. 4240; G. Szegö, *Math. Zeitschrift*, vol. 13 (1922), pág. 38. Ver también *Journal für die reine und angewandte Math.*, vol. 158 (1927), págs. 6-18. Ejemplo 16: *Jahresbericht der Deutschen Math. Vereinigung*, vol. 28 (1919), págs. 31-40; cf. pág. 38. Ejemplo 17: *Proceedings of the National Academy of Sciences*, vol. 33 (1947), págs. 218-221; cf. pág. 219. Ejemplo 19: *Journal für die reine und engewandte Math.*, vol. 151 (1921), páginas 1-31; ver teorema I, pág. 3.

** Cf. G. Polya y G. Szegö, *Analysis*, vol. 1, pág. VI.

$A_1^{(k)}, A_2^{(k)}, \ldots$ del cuadro, está relacionado con $R$, el número de raíces de la ecuación $f(x) = 0$ entre 0 y 1 (excluidos los extremos):

$$V(k) \geqq R,$$

y $V(k)$ sólo puede decrecer o permanecer invariable a medida que $k$ aumenta. ¿Alcanzará por fin $R$ la nunca creciente $V(k)$? Laguerre planteó la cuestión, pero la dejó sin resolver. M. Fekete probó que $V(k)$ alcanza, finalmente, $R$, dado que $R = 0$, y este caso particular sugirió la conjetura de que $V(k)$ alcanza siempre $R$. Yo tuve la suerte de observar otra conexión entre los números $A_n^{(k)}$ y la función $f(x)$: hay ciertos números positivos $B_n^{(k)}$ (independientes de la función $f(x)$) tales que

$$\lim_{k\to\infty} A_n^{(k)}/B_n^{(k)} = f(0) \text{ para } n \text{ fijo},$$

$$\lim_{n\to\infty} A_n^{(k)}/B_n^{(k)} = f(1) \text{ para } k \text{ fijo}.$$

Esto es, la dirección vertical en la ordenación cuadrada infinita se relaciona con $f(0)$, la dirección horizontal con $f(1)$. Estas dos direcciones extremas, la vertical y la horizontal, pueden recordar los *dos puntos extremos que piden ser unidos*: ¿qué pasa con las direcciones oblicuas intermedias? Por un momento tomemos la conjetura que tratamos de probar como garantizada: ésta implica alguna conexión entre las direcciones intermedias y los valores que $f(x)$ toma cuando $x$ varía entre 0 y 1. Esto sugiere que $A_n^{(k)}/B_n^{(k)}$ tiende a $f(x)$, donde $x$ está de algún modo en conexión con el límite de $n/k$. En efecto, al fin, encontré, y esto fue fácil de pro-bar, que

$$\lim_{k,\, n\to\infty} A_n^{(k)}/B_n^{(k)} = f(x), \text{ si } \lim n/(n+k) = x.$$

Esta relación da la clave de la solución al problema de Laguerre.

**13.** *Hay una línea recta con una dirección dada por un punto dado. Dibujar una paralela.* En conjunción con la serie de Fourier de una función positiva $f(x)$,

$$f(x) = a_0 + 2\sum_1^\infty (a_n \cos nx + b_n \operatorname{sen} nx),$$

O. Toeplitz consideró la ecuación en

$$(*) \quad \begin{vmatrix} a_0 - \lambda & a_1 - ib_1 & \ldots & a_{n-1} - ib_{n-1} \\ a_1 + ib_1 & a_0 - \lambda & \ldots & a_{n-2} - ib_{n-2} \\ & & \cdot & \\ a_{n-1} + ib_{n-1} & a_{n-2} + ib_{n-2} & \ldots & a_0 - \lambda \end{vmatrix} = 0.$$

Su investigación reveló que las $n$ raíces de esta ecuación

$$\lambda_{n1}, \quad \lambda_{n2}, \quad \lambda_{n3}, \quad \ldots \quad \lambda_{nn}$$

«imitan» los $n$ valores equidistantes de la función $f(x)$

$$f\left(\frac{2\pi}{n}\right), \quad f\left(\frac{4\pi}{n}\right), \quad f\left(\frac{6\pi}{n}\right), \quad \ldots \; f\left(\frac{2n\pi}{n}\right).$$

Por ejemplo, la media aritmética de las $n$ raíces

$$\frac{\lambda_{n1} + \lambda_{n2} + \ldots + \lambda_{nn}}{n} = a_0$$

corresponde a

$$\lim_{n\to\infty}\left[f\left(\frac{2\pi}{n}\right) + f\left(\frac{4\pi}{n}\right) + \ldots + f\left(\frac{2n\pi}{n}\right)\right]\frac{1}{n} = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx = a_0.$$

Intentemos *dibujar una paralela*: la media geométrica de $n$ raíces

$$[\lambda_{n1}\lambda_{n2}\ldots\lambda_{nn}]^{1/n} = D_n^{1/n}$$

donde $D_n$ denota la determinante con $n$ filas que obtenemos haciendo $\lambda = 0$ al lado izquierdo de la ecuación (*). Esto puede corresponder a

$$(**) \qquad \lim_{n\to\infty}\left[f\left(\frac{2\pi}{n}\right)f\left(\frac{4\pi}{n}\right) \ldots f\left(\frac{2n\pi}{n}\right)\right]^{1/n} = e^{\frac{1}{2\pi}\int_0^{2\pi}\log f(x)\,dx}$$

En este estadio resulta conveniente buscar un caso particular apropiado. Para la función particular

$$f(x) = a_0 + 2a_1\cos x + 2b_1 \operatorname{sen} x$$

no hay dificultad en calcular $D_n$ y $\lim_{n\to\infty} D_n^{1/n}$, y este límite llega a ser igual al valor (**): *una dirección general de investigación se encuentra con un caso particular apropiado*, y habría sido difícil no establecer la conjetura: para cualquier función positiva $f(x)$

$$\lim_{n\to\infty} D_n^{1/n} = e^{\frac{1}{2\pi}\int_0^{2\pi}\log f(x)\,dx}$$

**14.** *El caso más obvio puede ser el único posible.* Si los coeficientes $a_0, a_1, a_2, \ldots, a_n, \ldots$, de la serie de valores

$$a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots$$

son enteros, y una infinidad de estos enteros son diferentes de 0, la serie diverge obviamente en el punto $z = 1$, pues su término general no tiende a 0. En consecuencia, el radio de convergencia de tal serie de valores es $\leqq 1$. El valor extremo 1 del radio de convergencia puede ser alcanzado. Un ejemplo obvio es la serie

$$1 + z + z^2 + \ldots + z^n + \ldots;$$

representa una función de carácter analítico extremadamente sencillo, la función racional $1/(1 - z)$. Otro ejemplo es

$$z + z^2 + z^6 + \ldots + z^{n!} + \ldots;$$

que representa una función de carácter analítico extremadamente compleja, una función no continuable. (Este es el ejemplo más familiar de una serie para la que el círculo de convergencia es una línea singular.) Estos dos ejemplos obvios son de naturaleza opuesta. Sin embargo, cualquier serie de valor con coeficientes integrales y radio de convergencia 1, cuya naturaleza analítica puede ser averiguada, revela ser semejante a uno u otro de estos ejemplos opuestos: representa o una función racional o una no continuable. Los teoremas de E. Borel y P. Fatou mostraron que las clases extensivas de funciones de naturaleza intermedia no pueden ser representadas por una serie tal. Como el autor logró probar semejantes teoremas y extraer más ejemplos, descubrió suficiente evidencia para una conjetura: «Las funciones analíticas continuables *más obvias* representadas por una serie de valor de coeficientes integrales y el radio de convergencia 1, ciertas funciones racionales, son las *únicas funciones tales*». En otras palabras: «Si el radio de convergencia de una serie de valor con coeficientes integrales alcanza el valor extremo 1, la función representada es necesariamente de un carácter extremo: o es una función racional, extremamente sencilla, o es una función no continuable, extremamente complicada».

**15.** *Formando el modelo. El poder de las palabras.* Este ejemplo tiene varios aspectos, que trataré de examinar uno tras otro.

(1) E. Laguerre descubrió varias secuencias de número reales

$$\alpha_0, \alpha_1, \alpha_2, \ldots \alpha_n$$

con una curiosa propiedad: si la ecuación (desde otro punto arbitraria) de $n$ grados

$$\text{(I)} \qquad a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = 0$$

tiene sólo raíces reales, también la ecuación

$$\text{(II)} \qquad a_0\alpha_0 + a_1\alpha_1 x + a_2\alpha_2 x^2 + \ldots + a_n\alpha_n x^n = 0$$

(en la que se transforma la secuencia (I)) tendrá sólo raíces reales. Laguerre propuso, aunque lo dejó irresuelto, el siguiente problema: encontrar una condición necesaria y suficiente que caracterice esta clase de secuencias.

Es fácil encontrar condiciones necesarias: aplicamos una secuencia de la clase deseada a cualquier ecuación que se sabe que tiene sólo raíces reales, y luego obtenemos una ecuación transformada, cuyas raíces son todas necesariamente reales. Por ejemplo, aplicando la secuencia $\alpha_0, \alpha_1, \ldots, \alpha_n$ a las ecuaciones

$$1 - x^2 = 0, \quad x^2 - x^4 = 0, \quad x^4 - x^6 = 0, \quad \ldots$$

que, obviamente, tienen sólo raíces reales, encontramos que $\alpha_0$, $\alpha_2$, $\alpha_4$, $\alpha_6$, ..., son necesariamente del *mismo signo*, todos positivos o todos negativos (en el amplio sentido: 0 no es excluido). Aplicando la misma secuencia a las ecuaciones

$$x - x^3 = 0, \quad x^3 - x^5 = 0, \quad x^5 - x^7 = 0, \quad \ldots$$

encontramos que también $\alpha_1$, $\alpha_3$, $\alpha_5$, $\alpha_7$, ..., son necesariamente del mismo signo. Manteniendo estas observaciones en la cabeza y aplicando la secuencia a la ecuación

$$\text{(III)} \qquad 1 + \binom{n}{1} x + \binom{n}{2} x^2 + \ldots + x^n = 0$$

(cuyas $n$ raíces son todas iguales a $-1$) obtenemos la condición necesaria que las raíces de la ecuación

$$\text{(IV)} \qquad \alpha_0 + \binom{n}{1} \alpha_1 x + \binom{n}{2} \alpha_2 x^2 + \ldots + \alpha_n x^n + 0$$

son todas reales y del mismo signo.

Obtenemos esta última condición acumulando varias condiciones necesarias. Su acumulación será muy fuerte: *¿es bastante fuerte para formar una condición suficiente?* Si así fuese, la curiosa proposición se mantendrá: «Si ambas ecuaciones (I) y (IV) tienen sólo raíces reales, y las raíces de (IV) son todas del mismo signo, también la ecuación (II) tiene sólo raíces reales».

(2) Esta conjetura vino a mi mente bastante pronto, pero yo no

podía confiar en ella: parecía demasiado extraña. No obstante, en el caso $n = 2$ la conjetura fue fácilmente verificada (el caso $n = 1$ es totalmente trivial).

Sin embargo, por casualidad, me encontré con un teorema probado por E. Malo: Si la ecuación (I) tiene sólo raíces reales y la ecuación

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_n x^n = 0$$

tiene sólo raíces reales del mismo signo, la ecuación (II) debe tener sólo raíces reales. El teorema de Malo era muy análogo a esa extraña conjetura y la hacía mucho menos extraña. Además, el teorema de Malo era una consecuencia de esa conjetura, como podía verse fácilmente, y así, otra amplia e importante consecuencia de la conjetura había sido verificada: la conjetura se presentaba como mucho más fuerte.

(3) La conjetura puede ser nuevamente formulada de la siguiente manera: «Si la secuencia de números positivos $\alpha_0, \alpha_1, \ldots, \alpha_n$, transforma la ecuación (III) en otra con sólo raíces reales, trasformará una ecuación arbitraria con sólo raíces reales en una ecuación de la misma naturaleza». En otras palabras: la ecuación (III) *pone el modelo:* su respuesta a la secuencia $\alpha_0, \alpha_1, \ldots, \alpha_n$, es imitada por todas las ecuaciones cuyas raíces son todas reales.

(4) ¿Por qué la ecuación (III) pone el modelo? Porque todas sus raíces coinciden. Esta contestación es «correcta» en cierto sentido. De cualquier modo, como me di cuenta más tarde, las ecuaciones (o funciones) cuyas raíces (o ceros) coinciden juegan un papel análogo en varios problemas análogos: ponen el modelo (son *tonangebend*).

Sin embargo, como la conjetura en cuestión era todavía una conjetura, acudí a la siguiente «explicación»: Todas las raíces de la ecuación (III) son iguales a $-1$. Estas raíces están todas amontonadas en un punto del eje real, tan juntas como es posible. En tal situación son, comprensiblemente, las más inclinadas a saltar fuera del eje real. Por tanto, si la consecuencia $\alpha_0, \alpha_1, \ldots, \alpha_n$, aplicada al polinomio $(1 + x)^n$ con la mayoría de las raíces acumuladas no logra conducir estas raíces fuera del eje real, todavía tiene menos oportunidad de dispersar las raíces menos acumuladas de otros polinomios.

El valor lógico de esta explicación es obviamente nulo, pero ello no implica que su valor psicológico sea también nulo. Estoy convencido de que esta analogía fue muy importante para mí personalmente: me ayudó a mantener viva la conjetura por muchos años.

Mencionaré aquí que formulaciones verbales semejantes han sido a menudo conectadas con mi obra matemática. La frase al final del

ejemplo 14 es un ejemplo característico. Aquí tenemos dos ejemplos más.

Durante más de dos décadas estuve muy interesado en el conocido teorema de Fabry sobre series de valor. Hubo dos períodos: uno primero «contemplativo» y un segundo «activo». En este último trabajé con el teorema y encontré varias pruebas, extensiones y otros análogos a él. En el período contemplativo prácticamente no trabajé nada con el teorema, sólo lo admiré y lo recordaba de tiempo en tiempo bajo alguna formulación curiosa y estratégica, como la siguiente: «Si es infinitamente improbable que en una serie de valor un coeficiente elegido al azar sea diferente de 0, no es infinitamente probable, sino cierto, que la serie de valor no es continuable». Obviamente, esta sentencia no tiene ni lógica, ni mérito literario, sino que me sirvió personalmente para mantener vivo mi interés.

La idea de una cierta prueba se me ocurrió claramente, pero varios días después, en que no hacía nada para lograr la forma final de la prueba. Durante estos días estuve obsesionado por la palabra «trasplante». De hecho, esta palabra describe la idea decisiva de la prueba tan precisamente como es posible a una simple palabra describir una cosa tan compleja.

Di, por supuesto, varias explicaciones de este «poder de las palabras», pero quizá es mejor dejar las explicaciones hasta tener más ejemplos *.

**16.** *Esto es demasiado improbable para ser una mera coincidencia.* $f$ denota el número de factores primos de los enteros $n$, y llamemos $n$ «factorizado par» o «factorizado impar» según $f$ sea impar o par. Por ejemplo:

$$30 = 2 \times 3 \times 5 \text{ es factorizado impar.}$$

$$60 = 2 \times 2 \times 3 \times 5 \text{ es factorizado par.}$$

Números primos como 2, 3, 5, 7, 11, 13, 17, ... son factorizados impares, cuadrados como 4, 9, 16, ... son factorizados pares, y el número 1 puede ser considerado como factorizado par puesto que no tiene factores primos y 0 es un número par. Entre los doce primeros números

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| p | i | i | p | i | p | i | i | p | p | i | i |

cinco son factorizados par y siete impar.

---

* Cf. J. HADAMARD, *The psychology of invention in the mathematical field*, páginas 84-85.

Si miramos la sucesión de «pes» e «ies» en el esquema anterior, apepodemos detectar una sencilla regla. Las dos clases de números aparecen alternados irregularmente, de modo impredecible, *al azar*. La idea de casualidad se nos ocurre casi inevitablemente: es natural pensar que podemos obtener una secuencia semejante si, en vez de tomar el problema de factorizar un entero propuesto, echásemos una moneda al aire y escribiésemos «p» o «i» según salga cara o cruz. Es también natural sospechar que la moneda es «justa», que cara y cruz salen aproximadamente como una frecuencia igual, que las dos clases de enteros, los factorizados par e impar, son igualmente frecuentes.

Ahora podemos probar (la prueba es difícil) que entre los primeros $n$ enteros hay tantos factorizados par como impar siempre que $n$ sea largo. (El radio tiende a 1 como $n$ tiende a $\infty$.) Esto parece corroborar nuestra sospecha. Así podemos esperar con más confianza que los números factorizados par e impar se siguen unos a otros como una sucesión azarosa de caras y cruces. Esperándolo así, empecé a enumerar por cada $n$ qué clase de enteros, los factorizados en un sentido o en otro, están en mayoría entre los primeros $n$. Las primeras enumeraciones parecieron desequilibradas. Esto me sorprendió y procedí a alargar los valores de $n$: aún continuaron desequilibrados. Cuando llegué al $n = 1.500$ me cansé de cálculo numérico y admití que había suficiente evidencia experimental para admitir la conjetura: «Para $n \geqq 2$, los enteros factorizados par nunca están en mayoría entre los primeros $n$ enteros».

Lancemos una moneda, contemos cuántas caras y cuántas cruces salen en los primeros $n$ intentos. Puede suceder fácilmente que salgan menos caras que cruces. (Tomemos «menos» en el sentido amplio de «no más».) Podría también suceder que saliesen menos caras que cruces todas las veces a partir del segundo paso, para $n = 2, 3, 4, \ldots, 1.500$, pero esto no sucede fácilmente. *Esto es demasiado improbable para ser una mera coincidencia*, estamos tentados de decir. Sin embargo, lo improbable ha sucedido y lo hemos observado al jugar a cara y cruz con los enteros factorizados. *Tiene que haber alguna razón*. Entonces hice lo que había que hacer: tomé la conjetura como garantía (tentativamente, por supuesto) e intenté derivar consecuencias de ello. Tuve bastante suerte como para observar dos puntos. Primero, si la nueva conjetura fuese cierta, la verdad de una conjetura más lejana e importante debida a Riemann (sobre la función $\xi$) se seguiría necesariamente. Segundo, si fuese cierto algo más que la nueva conjetura (que los números factorizados par estuviesen definitivamente en minoría a partir de un cierto $n$) otra importante conjetura debida a Gauss (sobre la

clase de una forma cuadrática) se seguiría necesariamente. Ambos puntos parecen hablar de la nueva conjetura.

La conjetura no fue más allá, pero su suerte está todavía sin decidir. A. E. Ingham derivó consecuencias de ella que pueden hacerla menos digna de crédito. Por otro lado, D. H. Lehmer la verificó por cálculo numérico hasta $n = 600.000$.

**17.** *Perfeccionando la analogía.* «De todos los sólidos con un volumen dado la esfera tiene la superficie mínima.» Este es el teorema isoperimétrico clásico en el espacio; uno análogo a éste en física fue descubierto por H. Poincaré y estrictamante probado por G. Szegö: «De todos los sólidos con un volumen dado la esfera tiene la capacidad electrostática mínima». Es natural pensar que habrá más teoremas análogos, y yo estuve buscándolo. El campo de fuerza alrededor de un sólido cargado con electricidad es semejante al campo de flujo alrededor de un cuerpo sólido que se mueve con velocidad uniforme por un fluido ideal incomprensible. (Ambos campos no tienen origen y carecen de rotación, por lo que satisfacen la misma ecuación diferencial.) La capacidad del sólido en el campo electrostático corresponde groseramente a la «masa virtual» del sólido en el campo hidrodinámico. (El cuerpo que se mueve levanta el fluido y añade a su propia energía cinética la del fluido que se mueve; la «masa virtual» es un factor de esta energía cinética adicional.) Ambas, tanto la capacidad como la «masa virtual», están conectadas con la energía del campo correspondiente. Sin embargo, hay una notable diferencia: la capacidad depende sólo de la forma y el tamaño del sólido, pero la masa virtual depende también de la dirección del movimiento del sólido. Con objeto de hacer la analogía más perfecta construí un nuevo concepto: promediando la masa virtual en todas las direcciones posibles obtenemos la *masa virtual media*. Y así surge la conjetura: «De todos los sólidos con un volumen dado la esfera tiene la masa virtual media mínima».

El elipsoide es la única forma para la que la masa virtual ha sido explícitamente calculada en todas direcciones. En efecto, esto reveló que de todos los elipsoides con un volumen dado la esfera tiene la masa virtual media mínima: la conjetura se ha verificado en un caso particular importante. La conjetura podría también apoyarse por analogía: ha logrado probar la propiedad hidrodinámica mínima y análoga del círculo en dos dimensiones. De este modo apoyada la conjetura, merece ser públicamente formulada; al menos así lo creo.

La conjetura no ha sido aprobada ni desaprobada hasta la fecha, aunque G. Szegö y M. Schiffer han encontrado interesantes resultados conectados con ella que parecen apoyarla.

**18.** *Una nueva conjetura.* Consideremos el plano en que las coordenadas rectangulares son $x$ e $y$, y dentro de este plano una región $R$ rodeada por una curva cerrada $C$. Buscamos una función $u = u(x, y)$ que satisfaga una u otra de las condiciones límites.

$$(1) \quad u = 0,$$

$$(2) \quad \frac{\delta u}{\delta n} = 0,$$

a lo largo de la curva $C$ y la ecuación parcial diferencial

$$(3) \quad \frac{\delta^2 u}{\delta x^2} + \frac{\delta^2 u}{\delta y^2} + vu = 0$$

dentro de la región $R$. En (2) $n$ denota la normal a la curva $C$; en (3) $v$ denota una cierta constante. Así tenemos dos problemas diferentes: en el primero tenemos que resolver la ecuación diferencial (3) con la condición límite (1); en el segundo, con (2). Ambos problemas son importantes en física y están relacionados con varios problemas de vibración; los dos tienen la misma solución trivial: $u = 0$ idénticamente. Sólo para valores particulares de $v$ tiene el uno o el otro problema una solución no trivial, es decir, una solución $u$ que no desaparece idénticamente: el problema con la condición límite (1) para $v = \lambda_1, \lambda_2, \lambda_3, \ldots$, el problema con la condición límite (2) para $v = \mu_1, \mu_2, \mu_3, \ldots$ Esto es,

$$0 < \lambda_1 < \lambda_2 \leqq \lambda_3 \leqq \lambda_4 \leqq \ldots,$$

$$0 = \mu_1 < \mu_2 \leqq \mu_3 \leqq \mu_4 \leqq \ldots$$

Estos valores excepcionales $\lambda_1, \lambda_2, \lambda_3, \ldots$, y $\mu_1, \mu_2, \mu_3, \ldots$, reciben el nombre de valores propios del primero y segundo problemas, respectivamente. Los valores propios (o autovalores) están relacionados con las frecuencias de las vibraciones características en los correspondientes fenómenos físicos.

Deseamos establecer una nueva conjetura: *$A$ denota el área de la región $R$; luego, para $n = 1, 2, 3, \ldots$*

$$\mu_n < 4\pi n A^{-1} < \lambda_n.$$

Esta es una conjetura estimulante. Como la forma de la región $R$ puede variar arbitrariamente y $n$ se extiende a todos los enteros 1, 2, 3, ..., la conjetura cubre una inmensa variedad de casos particulares muchos de los cuales no son demasiado conocidos: la conjetura podría ser refutada por un resultado numérico concerniente a cualquiera de

estos casos particulares. Sin embargo, la conjetura no está demasiado mal apoyada.

(*a*) La conjetura es verificada para $n = 1, 2, 3, \ldots$, cuando $R$ es un rectángulo. Este fue, en efecto, el caso particular que sugirió la conjetura.

(*b*) La conjetura es verificada para $n = 1, 2, 3, \ldots$, cuando $R$ es de una forma arbitraria. Este caso particular es muy diferente del mencionado ne (*a*).

(*c*) La conjetura ha sido verificada por cálculo numérico en unos pocos casos particulares en que los *eigenvalues* pueden ser explícitamente calculados: para $n$ más allá de 25 y para formas especiales de $R$ (el círculo, unos pocos sectores circulares, unos pocos triángulos).

(*d*) Se desconoce desde hace tiempo que

$$\mu_1 \leqq \lambda_1, \quad \mu_2 \leqq \lambda_2, \quad \mu_3 \leqq \lambda_3, \ldots$$

Esto concuerda con la conjetura.

(*e*) Se conoce también que

$$\lim_{n\to\infty} \mu_n n^{-1} = \lim_{n\to\infty} \lambda_n n^{-1} = 4\pi A^{-1},$$

y esto también está de acuerdo con la conjetura.

Por supuesto, que ninguna de tales verificaciones pueden probar la conjetura u obligar a nadie a creer en ella en cualquier grado. Sin embargo, tales verficaciones acrecientan el interés de la conjetura, nos estimulan a ensayar nuevas consecuencias y añaden un aliciente a la investigación científica, que es la *raison d'être* de una conjetura.

Los detalles técnicos de la conjetura establecida (por alguno de los cuales estoy en deuda con Peter Szegö) serán publicados en otro sitio.

**19.** *Otra nueva conjetura.* «Si $F'(x)$ es una función algebraica y todos los coeficientes $a_1, a_2, a_1, \ldots$, de la serie

$$F(x) = a_1 x + a_3 x^2 + a_3 x^3 + \ldots$$

son enteros, también $F(x)$ es una función algebraica.» En otras palabras: «A menos que la integral de una función algebraica sea ella misma una función algebraica no puede ser representada por una serie de valores con coeficientes integrales».

Como ilustración consideremos el desarrollo

$$\text{arc sen } 2x = 2 \sum_{o}^{\infty} \binom{2n}{n} \frac{x^{2n+1}}{2n+1}.$$

La derivada de arc sen $2x$ es una función algebraica y todos los coeficientes del desarrollo de esta derivada son enteros, como hemos visto en la fórmula de arriba. Pero arc sen $2x$ no es una función algebraica. Por tanto, si la conjetura propuesta es verdadera, debe haber infinitos coeficientes en la extensión de arriba que no son enteros. Esto es fácil de verificar: !Si $2n + 1$ es un número primo, no es divisor de $2n!$ Así, el caso examinado apoya la conjetura, que es también apoyada (y ha sido sugerida) por el siguiente hecho: si sustituimos «función racional» por «función algebraica» en el propuesto enunciado conjetural, lo cambiamos por una verdad y un enunciado probado. La conjetura también es apoyada por analogías algo vagas, por la «atmósfera» que rodea el tema de las series de valor con coeficientes integrales; ver el ejemplo 14.

Hay muchos casos semejantes al del arc sen $2x$ que podrían ser fácilmente probados, pero que no han sido examinados aún. Sin una investigación de semejantes consecuencias accesibles tal conjetura no habría sido impresa. Aquí la he publicado más bien como un ejemplo de conjetura no desarrollada e imperfectamente apoyada.

**20.** *¿Qué es típico?* Hasta donde yo puedo ver, no hay nada en los últimos ejemplos (12 a 19) que esté en desacuerdo con la impresión general derivada de los ejemplos anteriores. En estos últimos ejemplos, como en los otros, la conjetura en cuestión estuvo doblemente apoyada: por unos hechos claros y por una «atmósfera general». La primera clase de apoyo, por hechos claros, me parece bien dentro de los patrones delineados en este capítulo y el precedente: la analogía y las consecuencias verificadas son prominentes, consecuencias meramente plausibles también juegan algún papel; el apoyo desde dos lados muy diferentes parece ser vital. En conjunto, los ejemplos 12 a 19 parecen ser típicos.

Ninguna de las conjeturas mencionadas en los ejemplos 13 a 17 han sido refutadas (hasta la fecha, al menos). Esto puede parecer atípico. No obstante, yo no he mencionado aquí otras muchas conjeturas mías que han sido refutadas a los pocos minutos, u horas, o días; tales conjeturas de poca vida son rápidamente olvidadas. Por supuesto, que sólo publico conjeturas que han superado todas las pruebas obvias que se me han ocurrido y han resistido el trabajo de algunos meses; éstas son las conjeturas más estimulantes, con mayor oportunidad de sobrevivir. La investigación consiste en formar muchas conjeturas, refutar la mayor parte de ellas y establecer unas cuantas.

# El azar, conjetura siempre presente

*... la probabilidad de que esta coincidencia sea una mera obra del azar es, por tanto, considerablemente menor a* $(1/2)^{60}$ *... De aquí que esta coincidencia debe ser generada por alguna causa, y se puede asignar a una causa que provea una perfecta explicación de los hechos observados.*—G. KIRCHHOFF *.

## 1. FENOMENOS AZAROSOS DE MASAS

En el uso cotidiano de la lengua usamos las palabras «probable», «verosímil», «plausible» y «digno de crédito» con significados que no son debidamente matizados. Ahora tomemos por separado la palabra «probable» y aprendamos a usarla en un significado específico, como término técnico de una rama de la ciencia que se llama «teoría de la probabilidad»**.

Esta teoría tiene una gran variedad de aplicaciones y aspectos, y, por tanto, puede ser presentada y concebida de varias maneras. Algunos autores la consideran como una teoría puramente matemática; otros, como una clase o rama de la lógica y, aun otros, como una parte del estudio de la naturaleza. Estos puntos de vista pueden o no ser incompatibles. Tenemos que empezar a estudiar uno de ellos, pero sin entregarnos a ninguno. En el próximo capítulo cambiaremos en algo nuestra posición, pero en éste elegimos el punto de vista que es más conveniente en la mayor parte de las aplicaciones y que el principiante puede dominar más rápidamente. Nosotros aquí consideramos la teoría de la probabilidad como una parte del estudio de la naturaleza, como la teoría de ciertos fenómenos observables, los *fenómenos azarosos de masas* ***. Es fácil

* *Abhandlungen der k. Akademie der Wissenschaften*, Berlín, 1861, págs. 78-80.

** En lo anterior, las palabras «probable» y «probabilidad» han sido usadas a veces en un sentido no técnico, pero esto será cuidadosamente evitado en el capítulo presente y en el próximo. Las palabras «verosímil» y «verosimilitud» serán presentadas como términos técnicos al final de este capítulo.

*** En este punto esencial, y en otros varios, la presente exposición sigue los puntos de vista de RICHARD VON MISES, aunque se desvía de su definición de probabili- matemática; cf. su libro *Probability, Statistics and Truth.*

comprender con claridad lo que este término significa si comparamos unos cuantos ejemplos familiares de tales fenómenos.

(1) *Chaparrón.* El chaparrón es un fenómeno de masas. Consiste en un número muy grande de sucesos singulares, en este caso la caída de un gran número de gotas de lluvia. Estas gotas, aunque muy semejantes unas a otras, difieren en varios aspectos: en el tamaño, en el lugar donde caen, etc. Hay algo en la conducta de las gotas de lluvia que puede ser considerado como «azar». Al objeto de comprender claramente el significado de este término imaginemos un experimento.

Observemos las primeras gotas sobre el pavimento cuando la lluvia empieza a caer. Miremos el pavimento en mitad de una gran plaza pública, suficientemente lejos de edificios, árboles o cualquier otra cosa que pueda obstruir la lluvia. Fijemos nuestra atención sobre dos piedras a las que podemos llamar «piedra del lado derecho» y «piedra del lado izquierdo». Veamos las gotas que caen sobre estas piedras y el orden en que las golpean. La primera gota cae sobre la piedra izquierda, la segunda sobre la derecha, la tercera de nuevo sobre la derecha, la cuarta sobre la izquierda, y así sucesivamente sin aparente regularidad como

*I D D I I I D I D I D D I D D*

(*I* para izquierda, *D* para derecha). No hay regularidad en la sucesión de gotas de lluvia. En efecto, habiendo observado un cierto número de gotas no podemos predecir qué camino seguirá la próxima gota. Hemos anotado arriba quince frecuencias. Mirando éstas ¿podemos predecir si la próxima gota irá a la *I* o a la *D*? Obviamente, no. Por otro lado, hay una cierta regularidad en la sucesión de las gotas. Así, podemos predecir confiadamente que al final de la lluvia las dos piedras estarán igualmente húmedas. Es decir, el número de gotas que caen en cada piedra será casi proporcional al área de su superficie horizontal libre. Nadie duda que esto es así, y los meteorologistas lo asumen al construir sus pluviómetros. Pero aquí hay algo paradójico. Podemos prever lo que sucederá a largo alcance, pero no los detalles. El chaparrón es un típico fenómeno de masas fortuito, *impredecible en ciertos detalles, predecible en ciertas proporciones numéricas del conjunto.*

(2) *Niños entre los recién nacidos.* En un hospital los bebés recién nacidos son registrados en el orden de su nacimiento. Niños y niñas (*M* y *H*, respectivamente) se siguen unos a otros sin aparente regularidad como

*H M M H M H H M M H H M M M H H,*

Aunque no predecimos los detalles de esta sucesión azarosa, podemos muy bien predecir un importante rasgo del resultado final obtenido por la suma de todos los registros durante un año en los Estados Unidos: el número de chicos será mayor que el de chicas y, de hecho, la razón de estos dos números será muy poco diferente de la razón 51,5 : 48,5, El número de nacimientos en los Estados Unidos es de unos tres millones por año. Tenemos aquí de nuevo un azaroso fenómeno de masas de considerables dimensiones.

(3) *Un juego de azar.* Lancemos una moneda al aire repetidas veces, observando del lado que cae cada vez, «cara» o «cruz», o, con otras palabras, «anverso» y «reverso» ($A$ y $R$). Así obtenemos una sucesión sin regularidad aparente como

$$R\,A\,A\,A\,R\,A\,R\,A\,A\,R\,A\,R\,A\,R\,R.$$

Si tenemos la paciencia da lanzar la moneda unos cuantos cientos de veces, aparece una definida razón de anversos y reversos, que no cambia mucho si prolongamos el experimento aún más. Si nuestra moneda es «justa», la razón 50 : 50 de caras a cruces aparece a la larga. Si la moneda es «injusta» (por estar torcida, por ejemplo), alguna otra razón aparecerá a la vista. En cualquier caso, de nuevo aparecerán los rasgos de un azaroso fenómeno de masas. A la larga siempre surgen proporciones constantes, aunque los detalles son impredecibles. Hay una cierta regularidad en el conjunto, a despecho de la irregularidad en los casos individuales.

## 2. El concepto de probabilidad

En el año 1943 el número de nacimientos en los Estados Unidos, machos, hembras y total, fue de

1.506.959 1.427.901 2.934.860,

respectivamente. Llamamos

1.506.959 la frecuencia del nacimiento de machos,

1.427.901 la frecuencia del nacimiento de hembras.

Llamamos

$$\frac{1.506.959}{2.934.860} = 0{,}5135$$

la frecuencia relativa del nacimiento de machos y

$$\frac{1.427.901}{2.934.860} = 0{,}4865$$

la frecuencia relativa del nacimiento de hembras. En general, si un suceso de cierta clase ocurre en $m$ casos de $n$, llamamos $m$ la *frecuencia* de ocurrencia de esta clase de sucesos y $m/n$ su *frecuencia relativa.*

Imaginemos que, durante todo el año, los nacimientos son sucesivamente registrados en todos los Estados Unidos (como en el hospital que hemos mencionado en la sección anterior). Si miramos la sucesión de nacimientos de machos y hembras tendremos ante nosotros una serie extremadamente larga de casi tres millones de entradas, que empiezan como

*H M M H M H M M H H M M M H H .*

Según el fenómeno de masas se desenvuelve, tenemos, en cada estadio de observación, una cierta frecuencia de nacimientos machos, igual que una cierta frecuencia relativa. Notemos, tras las observaciones 1, 2, 3, ..., las frecuencias y las frecuencias relativas encontradas hasta ese punto:

| Observaciones | Suceso | Frecuencia de M | Frecuencia relativa |
|---|---|---|---|
| 1 | *H* | 0 | 0/1 = 0,000 |
| 2 | *M* | 1 | 1/2 = 0,500 |
| 3 | *M* | 2 | 2/3 = 0,667 |
| 4 | *H* | 2 | 2/4 = 0,500 |
| 5 | *M* | 3 | 3/5 = 0,600 |
| 6 | *H* | 3 | 3/6 = 0,500 |
| 7 | *M* | 4 | 4/7 = 0,571 |
| 8 | *M* | 5 | 5/8 = 0,625 |
| 9 | *H* | 5 | 5/9 = 0,556 |
| 10 | *H* | 5 | 5/10 = 0,500 |
| 11 | *M* | 6 | 6/11 = 0,545 |
| 12 | *M* | 7 | 7/12 = 0,583 |
| 13 | *M* | 8 | 8/13 = 0,615 |
| 14 | *H* | 8 | 8/14 = 0,571 |
| 15 | *H* | 8 | 8/15 = 0,533 |

Hasta donde hemos tabulado, la frecuencia relativa oscila bastante fuertemente (entre los límites 0,000 y 0,667). Pero aquí tenemos un

número de observaciones muy pequeño. A medida que tenemos más y más, las oscilaciones de la frecuencia relativa se van haciendo menos y menos violentas, y podemos confiar en que al final oscilará muy poco en torno al valor definitivo de 0,5135. *Cuando el número de observaciones aumenta, la frecuencia relativa aparece fijada a un valor estable, a despecho de las irregularidades impredecibles de detelle.* Tal conducta, la aparición de una frecuencia relativa estable a largo plazo, es típica de los fenómenos de masas.

Un objetivo importante de cualquier teoría de tales fenómenos debe ser predecir la final y estable frecuencia relativa o la *frecuencia relativa a largo alcance.* Consideraremos el *valor teórico de la frecuencia relativa a largo alcance* y llamaremos a este valor teórico *probabilidad.*

Tenemos que aclarar este concepto de probabilidad. Naturalmente, empezamos con el estudio de los fenómenos de masas cuya frecuencia relativa a largo alcance podemos predecir con un grado de confianza razonable.

(1) *Bolas en un saco.* Un saco contiene $p$ bolas de varios colores entre las que hay exactamente $f$ bolas blancas. Utilicemos esto para producir un fenómeno de masas. Extraigamos una bola, miremos su color y escribamos $B$ si la bola es blanca o $D$ si es de color diferente. Metamos la bola pintada dentro del saco y luego la mezclamos dentro con las otras; después extraemos otra bola y apuntamos su color, $B$ o $D$. Al proceder así obtenemos una secuencia azarosa semejante a la considerada en la sección 1:

$$B\,D\,D\,D\,B\,D\,D\,B\,B\,D\,D\,D\,B\,B\,D\,.$$

¿Cuál es la frecuencia relativa a largo alcance de las bolas blancas?

Examinemos las circunstancias en que podemos predecir la frecuencia deseada con confianza razonable. Asumimos que las bolas son homogéneas y perfectamente esféricas, hechas del mismo material y teniendo el mismo radio. Sus superficies son igualmente tersas y su diferente coloración influye imperceptiblemente sobre su comportamiento mecánico, si es que influye algo. La persona que extrae las bolas está ciega o se ha procurado de algún modo que no vea las bolas. La posición de las bolas en el saco varía de unas a otras, es impredecible, está más allá de nuestro control. Pero las circunstancias permanentes están todas bajo control: tienen todas la misma forma, tamaño y peso, son *indistinguibles* por la persona que las extrae.

Bajo tales circunstancias no vemos razón para preferir una bola a otra y, naturalmente, esperamos que, a largo plazo, cada bola será

extraída *igualmente a menudo*. Supongamos que tenemos paciencia para extraer 10.000 bolas. Entonces podemos esperar que cada una de las $p$ bolas aparecerá unas

$$\frac{10.000}{p} \text{ veces.}$$

Hay $f$ bolas blancas. Por tanto, en 10.000 extracciones, esperamos alcanzar blancas

$$f\,\frac{10.000}{p} = 10.000\,\frac{f}{p} \text{ veces;}$$

ésta es la frecuencia esperada de bolas blancas. Para obtener la frecuencia relativa dividiremos la frecuencia por el número de observaciones o, en este caso, bolas extraídas, es decir, por 10.000. Y así nos vemos conducidos al siguiente enunciado: la frecuencia relativa a largo alcance, o *probabilidad*, de las bolas blancas es $f/p$.

Las letras $f$ y $p$ están elegidas conforme al modo tradicional de expresión. Cuando extraemos una de las $p$ bolas elegimos uno de los $p$ *posibles* casos. Tenemos buenas razones (en condiciones iguales de las $p$ bolas) para no preferir ninguno de los $p$ posibles casos a cualquier otro. Si deseamos que sea extraída una bola blanca (por ejemplo, cuando apostamos por el blanco), las $f$ bolas blancas se nos presentarán como casos *favorables*. De aquí que podamos describir la probabilidad $f/p$ como la *razón del número de casos favorables respecto al número de casos posibles*.

Sacar una bola del saco, meter luego la bola dentro del saco, menearle, sacar otra bola, y repetir esto $n$ veces parece ser una ocupación bastante tonta. ¿Perderemos nuestro tiempo estudiando un juego tan primitivo? No lo creo. El saco y las bolas, manejados de la manera descrita, engendran un azaroso fenómeno de masas que resulta particularmente sencillo y accesible. La generalización empieza por los casos más sencillos, los casos particulares más transparentes. La dinámica nació como ciencia el día que Galileo empezó a estudiar la caída de los cuerpos pesados. La ciencia de la probabilidad nació cuando Fermat y Pascal empezaron a estudiar los juegos de azar que dependen de echar los dados o de sacar una carta de la baraja o una bola de un saco. Las leyes y conceptos fundamentales de la dinámica pueden ser extraídos a partir del sencillo fenómeno de la caída de los cuerpos. Nosotros utilizamos el saco y las bolas para comprender el concepto fundamental de probabilidad.

(2) *Chaparrón.* Ahora volvemos a considerar el azaroso fenómeno de masas que iniciamos en la sección 1. El área de una superficie hori-

zontal es $P$ y el área de una cierta porción de esta superficie es $F$; ver la figura 14.1. Observemos las gotas de lluvia caídas en el área $P$ e interesémonos en la frecuencia de las gotas de lluvia caídas en el subárea $F$. Nos sentimos inclinados a predecir sin duda la frecuencia relativa a largo alcance: la fracción de la lluvia total sobre el área que cae en el subárea será muy aproximada a $F/P$ siempre que la lluvia conste de más de unas cuantas gotas. En otras palabras: la probabilidad de que una gota de lluvia que cae sobre el área $P$ caiga dentro de la porción de área $F$ es de $F/P$. Si idealizamos el chaparrón y consideramos una gota de lluvia como un punto geométrico podemos decir también: la probabilidad de que un punto que cae en el área $P$ caiga dentro del subárea $F$ es $F/P$.

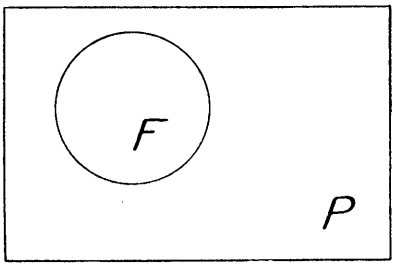



Fig. 14.1. La probabilidad definida por el chaparrón.

En el último enunciado consideramos cada punto del área $P$ como un caso posible y cada punto del subárea $F$ como un caso favorable. El número de casos favorables como el de casos posibles es infinito, y no tendría sentido hablar de la razón de números infinitos. No obstante, podemos considerar el área de una superficie como la *medida* de los puntos contenidos en la superficie. Usando este término podemos describir la probabilidad $F/P$ como la *razón de la medida de los casos favorables respecto a la medida de los casos posibles.*

## 3. Usando el saco y las bolas

Al derivar el principio fundamental de la estadística, Lagrange remplazó un sistema arbitrario de fuerzas por un conveniente sistema de poleas. A la luz de este argumento (cuyos detalles no son aquí necesarios *) cualquier caso de equilibrio aparece como una apropiada combinación de poleas correctamente balanceadas. El cálculo de probabilidad puede ser visto de una manera semejante; de hecho, tal visión es sugerida en los primeros tiempos de esta ciencia. Desde este punto de vista,

* Ver E. Mach, *Die Mechanik*, págs. 59-62.

cualquier problema de probabilidad parece comparable a un problema adecuado de sacos que contienen bolas, y cualquier azaroso fenómeno de masas aparece semejante, en ciertos aspectos esenciales, a sucesivas extracciones de bolas a partir de un sistema de sacos convenientemente combinados. Ilustremos esto con unos cuantos ejemplos sencillos.

(1) En lugar de lanzar al aire una moneda «justa» podemos sacar una bola de un saco que contiene sólo dos bolas, una de ellas marcada con una *A* y la otra con una *R* (anverso y reverso). En lugar de echar un dado sin deformar podemos sacar una bola de un saco que contiene seis bolas, señaladas con 1, 2, 3, 4, 5 y 6 puntos, respectivamente. En lugar de sacar una carta de una baraja podemos sacar una bola de un saco que contiene cincuenta y dos bolas*, adecuadamente marcadas. Parece ser claro intuitivamente que al sustituir un saco con bolas por monedas, dados, cartas y otros ingenios semejantes de un modo apropiado no cambiamos las ventajas de los juegos de azar corrientes. Al menos no cambiamos las oportunidades en las versiones idealizadas de estos juegos en que los medios usados (monedas, dados, etc.) se supone que son perfectamente simétricos y, por tanto, ciertas oportunidades fundamentales perfectamente iguales.

(2) Para estudiar el azar en la distribución de niños y niñas entre los recién nacidos podemos sustituir el fenómeno de masas por sucesivas extracciones de un saco que contiene 1.000 bolas, 515 marcadas con *M* y 485 con *H*. Esta sustitución es, por supuesto, teórica y, como toda teoría, está destinada a ser tentativa y aproximativa. Sin embargo, la cuestión es que el saco y las bolas nos incapacitan para formular una teoría.

(3) Un meteorologista registra la sucesión de días con lluvia y sin lluvia en una cierta localidad. Sus observaciones parecen mostrarnos que, en conjunto, cada día tiende a parecerse al día anterior: a días sin lluvia parecen seguir días sin lluvia más fácilmente que días con lluvia y, de modo semejante, a días lluviosos parecen seguir días lluviosos más fácilmente que días sin lluvia. Por supuesto, que una confiable regularidad aparece sólo en una larga serie de observaciones; los detalles son irregulares, parecen ser azarosos.

El meteorologista puede desear expresar sus impresiones de modo más claro que el que hemos bosquejado aquí. Si deseamos formular una teoría en términos de probabilidad podemos considerar tres sacos. Cada saco contiene el mismo número de bolas, digamos 1.000. Algunas de estas bolas son blancas, las otras son negras (blancas para los días sin

* Hay que tener en cuenta que la baraja inglesa tiene cincuenta y dos cartas a diferencia de la española en que suelen usurse sólo cuarenta. *(N. del T.)*

lluvias, negras para los lluviosos). Sin embargo, hay una diferencia importante entre los sacos. Cada saco lleva una inscripción fácilmente visible a las personas que sacan las bolas. Un saco lleva la inscripción «COMIENZO», otro «TRAS BLANCO» y el tercero «TRAS NEGRO». La razón de las bolas de diferente color es diferente en los distintos sacos. En cada saco la razón de las bolas blancas respecto a las bolas negras se aproxima a la razón observable de los días sin lluvia respecto de los días lluviosos, aunque en circunstancias diferentes. En el saco «COMIENZO» la razón es la de los días sin lluvia con relación a los días lluviosos durante todo el año; en el saco «TRAS BLANCO» la razón es la de los días sin lluvia respecto a los días lluviosos que siguen a un día sin lluvia; en el saco «TRAS NEGRO», la de los días sin lluvia respecto a los días lluviosos que siguen a un día lluvioso. Por tanto, el saco «TRAS BLANCO» contiene más bolas blancas que el saco «TRAS NEGRO». Las bolas son extraídas sucesivamente y cada bola extraída, una vez que hemos tomado nota de su color, es remplazada dentro del saco del que la hemos sacado. El saco «COMIENZO» es usado sólo una vez para la primera bola. Si la primera bola es blanca, utilizaremos el saco «TRAS BLANCO» para la segunda bola; pero si la primera bola es negra, la segunda bola es sacada del saco «TRAS NEGRO». Y así sucesivamente, el color de la bola sacada determina el saco del que la próxima bola será sacada.

Es sólo una teoría que la sucesión de bolas blancas y negras extraídas en las circunstancias descritas imita la sucesión de los días sin y con lluvia de un modo razonablemente aproximado. Pero, según se ve, esta teoría no está fuera de lugar. En cualquier caso, esta teoría, u otra semejante, merecería ser confrontada con las observaciones.

(4) Tomemos un texto inglés (de Shakespeare, si preferimos) y remplacemos cada una de las letras *a*, *e*, *i*, *o*, *u* e *y* por *V* y cada una de las veinte letras restantes por *C*. (*V* significa vocal y *C* significa consonante.) Así obtendremos un patrón como

$$C\,V\,C\,V\,V\,C\,C\,V\,C\,C\,V\,C\,V\,C\,C.$$

Esta secuencia irregular es de algún modo opuesto a la que examinamos en la subsección anterior (3): cada día tiende a ser igual al día anterior, pero cada letra tiende a ser desigual a la letra anterior. Aun así, podríamos imitar la sucesión de bolas blancas y negras sacadas de tres sacos que llevan las mismas inscripciones de antes (en la subsección (3)), aunque la razón de las bolas blancas con respecto a las negras no puede ser la misma de antes. Para imitar de modo realista la sucesión de vocales y consonantes el saco «TRAS BLANCO» deberá contener *menos* bolas blancas que el saco «TRAS NEGRO».

(5) Tenemos dos sacos. El primero contiene $p$ bolas entre las que $f$ son de color blanco. El segundo saco contiene $P$ fichas entre las que $F$ son blancas. Usando ambas manos, sacamos de los dos sacos al mismo tiempo una bola con la mano izquierda y una ficha con la derecha. ¿Cuál es la probabilidad de que ambas, bola y ficha, sean blancas?

Sin duda, podríamos repetir este experimento lo bastante a menudo, quizá unos miles de veces, para obtener un valor aproximado de la probabilidad deseada. Sin embargo, podemos también intentar intuirlo, y esto es mucho más interesante.

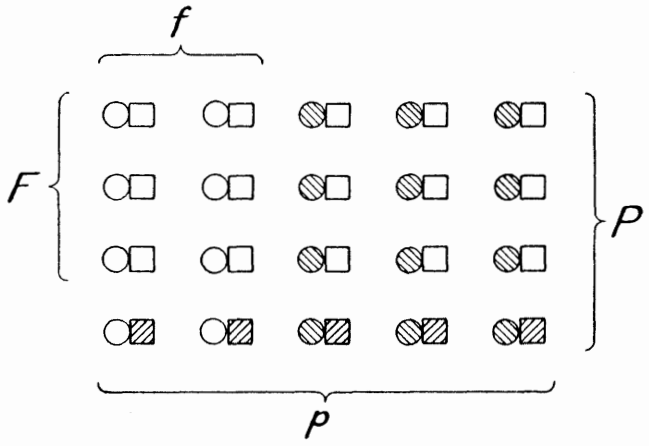



Fig. 14.2. Sucesos independientes.

El resultado de las dos extracciones simultáneas es una «pareja», que consta de una bola y una ficha. Hay $p$ bolas y $P$ fichas. Como una bola se apareja con una ficha, tenemos $pP$ posibles parejas; éstas aparecen en la figura 14.2, donde $p = 5, f = 2\ P = 4,\ F = 3$. No hay razón para preferir una de las $p$ bolas a cualquiera otra, o una de las $P$ fichas a cualquiera de éstas. Por lo mismo, *no parece haber razón para preferir una de las pP parejas a cualquiera otra.* En efecto, al realizar el experimento con los dos sacos yo suponía proceder con los ojos vendados, al azar, de modo que cada mano sacase independientemente de la otra. «Que tu mano izquierda no sepa lo que hace tu mano derecha.» Resulta increíble que la suerte de la bola que saco con la mano izquierda sea influida por la ficha que saco con la derecha. ¿Por qué la ficha número 1 atrae más a la bola número 1 que la ficha número 2?

Y así, podemos imaginar un saco que contiene $pP$ objetos mecánicamente indistinguibles (cada objeto es una pareja, una bola ligada a una ficha); una extracción de este saco resulta *equivalente* a las extrac-

ciones simultáneas de los dos sacos descritos al comienzo. Tenemos, pues, $pP$ casos posibles; nos falta encontrar el número de casos favorables. Una ojeada a la figura 14.2 nos muestra que hay $fF$ parejas que constan de una bola blanca y una ficha del mismo color. De este modo obtenemos el valor de la probabilidad deseada: ésta es

$$\frac{fF}{pP} = \frac{f}{p} \cdot \frac{F}{P}$$

el *producto* de dos probabilidades. En efecto, $f/p$ es la probabilidad de extraer una bola blanca del primer saco y $F/P$ la probabilidad de extraer una ficha blanca del segundo saco.

El punto esencial en el caso de las bolas y las fichas es que la extracción de una no influye en la suerte de la otra. En la terminología usual del cálculo de probabilidad tales sucesos reciben el nombre de *independientes* el uno del otro; la unión de ambos sucesos es considerada como un suceso *compuesto*. La consideración precedente motiva la regla: *La probabilidad de un suceso compuesto es el producto de las probabilidades de los sucesos constituyentes, habida cuenta que estos sucesos constituyentes son mutuamente independientes.*

## 4. El cálculo de probabilidad. Hipótesis estadística

La teoría de la probabilidad es, como hemos visto, una parte del estudio de la naturaleza, la teoría de los fenómenos azarosos de masas.

El logro más sorprendente de las ciencias físicas es la predicción. Los astrónomos predicen con precisión los eclipses de Sol y de Luna, la posición de los planetas y el retorno de los planetas que escapan por varios años a la observación. Un gran astrónomo (Leverrier) logró también predecir la posición de un planeta (Neptuno), cuya existencia se desconocía anteriormente. La teoría de la probabilidad predice las frecuencias de ciertos fenómenos de masas con algún éxito.

Los astrónomos basan sus predicciones sobre las primeras observaciones, las leyes de la mecánica, la ley de gravitación y una serie de largos y difíciles cálculos. Cualquier rama de la ciencia física basa sus predicciones en alguna teoría o, podemos decir, en alguna conjetura, puesto que no hay una teoría cierta y todas las teorías no son más que conjeturas más o menos razonables, mejor o peor sustentadas. Al intentar predecir las frecuencias de un cierto fenómeno de masas a partir de la teoría de la probabilidad no tenemos más remedio que hacer una cierta

asunción teórica del fenómeno. Tal asunción, que ha de ser expresada en términos de probabilidad, recibe el nombre de *hipótesis estadística.*

Cuando aplicamos la teoría de la probabilidad tenemos que calcular probabilidades (que son valores teóricos y aproximados de frecuencias relativas). Cuando intentamos encontrar una probabilidad tenemos por delante un problema que resolver. La incógnita de este problema es la probabilidad deseada. Mas, para determinar esta incógnita, necesitamos los datos y las condiciones de nuestro problema. Los datos son usualmente probabilidades, y las condiciones, que dependen de la relación de la probabilidad desconocida a la probabilidad dada, constituyen una hipótesis estadística.

Como en las aplicaciones de la teoría de la probabilidad el cálculo de probabilidades juega un papel prominente, esta teoría recibe usualmente el nombre de *cálculo de probabilidad.* Así, el objetivo del cálculo de probabilidad es calcular nuevas probabilidades sobre la base de probabilidades dadas y las hipótesis estadísticas dadas.

El lector que desee continuar con la parte restante de este capítulo debe, o bien conocer los elementos del cálculo de probabilidad, o bien dar por ciertos algunos resultados derivados de estos elementos. La mayoría de las veces el texto establecerá los resultados sin derivación; las derivaciones serán dadas después en la primera parte de los ejercicios y comentarios que siguen a este capítulo y en las soluciones correspondientes. Aun con todo, si el lector no comprueba la derivación de los resultados debía echar una mirada a las asunciones teóricas subyacentes. Podemos hacer tales asunciones comprensibles intuitivamente: comparar los azarosos fenómenos de masas que examinemos a extracciones de sacos adecuadamente llenos bajo apropiadas condiciones, como en la anterior sección 3.

Las aplicaciones del cálculo de probabilidad tienen interminables variedades. Las secciones siguientes de este capítulo intentan ilustrar los tipos principales de aplicaciones por medio de ejemplos elementales adecuados. La atención será puesta sobre la motivación de estas aplicaciones, es decir, sobre las consideraciones preliminares que nos conducen a elegir el razonamiento plausible.

## 5. Predicción correcta de frecuencias

En los comienzos de su historia el cálculo de probabilidad fue esencialmente una teoría de ciertos juegos de azar. Pero las predicciones de esta teoría no han sido comprobadas experimentalmente a larga escala

hasta tiempos modernos. Empecemos por examinar un experimento de esta clase.

(1) W. F. R. Weldon echó doce dados 26.306 veces, observando cada vez cuántos de estos doce dados exhiben más de cuatro puntos *. Los resultados de sus observaciones son señalados en la columna (4) de la tabla I; la columna (1) muestra el número de dados, entre los doce, que han sacado cinco o seis puntos. Así, en los 26.306 intentos nunca sucedió que todos (cada uno de los doce) mostraran más de cuatro puntos. El caso más frecuente fue aquel en que cuatro de los doce dados mostraron cinco o seis puntos; esto ocurrió 6.114 veces.

TABLA I

| (1) Número de 5 ó 6 | (2) Exceso I | (3) Predicho I | (4) Observado | (5) Predicho II | (6) Exceso II |
|---|---|---|---|---|---|
| 0 | + 18 | 203 | 185 | 187 | + 2 |
| 1 | + 67 | 1216 | 1149 | 1146 | — 3 |
| 2 | + 80 | 3345 | 3265 | 3215 | — 50 |
| 3 | + 101 | 5576 | 5475 | 5465 | — 10 |
| 4 | + 159 | 6273 | 6114 | 6269 | + 155 |
| 5 | — 176 | 5018 | 5194 | 5115 | — 79 |
| 6 | — 140 | 2927 | 3067 | 3043 | — 24 |
| 7 | — 76 | 1255 | 1331 | 1330 | — 1 |
| 8 | — 11 | 392 | 403 | 424 | + 21 |
| 9 | — 18 | 87 | 105 | 96 | — 9 |
| 10 | — 1 | 13 | 14 | 15 | + 1 |
| 11 | — 3 | 1 | 4 | 1 | — 3 |
| 12 | 0 | 0 | 0 | 0 | 0 |
| Total | 0 | 26.306 | 26.306 | 26.306 | 0 |

¿Cómo puede la teoría predecir los números observados que encontrados en la columna (4) de la tabla I? Si asumimos que los dados son «justos» y que los intentos con dados diferentes, o con los mismos dados en veces diferentes, son independientes unos de otros, podemos calcular las probabilidades relevantes. Con esta asunción (que es propiamente llamada una «hipótesis estadística») la probabilidad de que exactamente cuatro dados de doce exhiban cinco o seis puntos es

$$P = 495\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^8 = \frac{126.720}{531.441}.$$

* *Philosophical Magazine*, serie 5, vol. 50, 1900, págs. 167-169; en un artículo de KARL PEARSON.

Ahora, por definición, la probabilidad es el valor teórico de la frecuencia relativa de largo alcance. Si el suceso con probabilidad $P$ se muestra $m$ veces en $n$ intentos, esperamos que

$$\frac{m}{n} = P, \text{ aproximadamente,}$$

o

$$m = Pn, \text{ aproximadamente.}$$

Por tanto, esperaremos que exactamente cuatro dados exhibirán cinco o seis puntos de los doce dados volteados en, aproximadamente,

$$Pn = \frac{126.720}{531.441}\ 26.306 = 6.273$$

casos de $n = 26.306$ intentos. (Observemos que podemos calcular este número 6.273 antes de empezar el ensayo.) Ahora bien, este valor predicho 6.273 no parece ser «muy diferente» del número observado 6.114, y así, nuestra primera impresión sobre la aplicabilidad práctica de la teoría de la probabilidad puede ser bastante buena.

El número 6.273 aparece en la columna (3) de la tabla I en el lugar adecuado, en la misma fila que el número 4 de la columna (1). Todos los números de la columna (3) son calculados de modo semejante. Al objeto de comparar más convenientemente los valores predichos en la columna (3) con los números observados de la columna (4) escribimos las diferencias (predichos menos observados) en la columna (2). Teniendo en cuenta su significado examinemos las columnas (2), (3) y (4). ¿Es satisfactorio el acuerdo entre experiencia y teoría? ¿Son los números observados suficientemente próximos a los valores predichos?

Hay, evidentemente, algún acuerdo entre las columnas (3) y (4). Ambas columnas tienen el mismo aspecto general: el máximo es alcanzado en el mismo punto (en la misma fila) y los números van aumentando primero hasta el máximo y luego disminuyen continuamente hasta 0 del mismo modo en ambas columnas. La desviación de los números observados respecto de los valores predichos parece ser relativamente pequeña en la mayoría de los casos; el acuerdo, a una primera ojeada, parece bastante bueno. Por otro lado, el número de intentos, 26.306, es muy largo. ¿Son las desviaciones suficientemente pequeñas a la vista de tan gran número de intentos?

Esta parece ser la pregunta correcta. Sin embargo, la contestación está fuera de alcance; debemos posponerla hasta saber un poco más;

ver la sección 7 (3). Aun así, sin ningún conocimiento especial, sólo con un poco de sentido común, podemos sacar una conclusión aguda de la tabla I. Un físico notaría en seguida el siguiente punto de las columnas (3) y (4). Las diferencias están apuntadas en la columna (2): algunas de éstas son positivas; otras, negativas. Si estas diferencias estuviesen distribuidas por azar, los signos + y — estarían entremezclados de modo desordenado. Sin embargo, los signos + y — están claramente separados: los valores teóricos son demasiado grandes hasta cierto punto y demasiado pequeños a partir de ese punto. En tal caso, los físicos hablan de una desviación *sistemática* de la teoría respecto de la experiencia, y consideran tal desviación sistemática como una grave objeción contra la teoría.

Y así, el acuerdo entre la teoría de la probabilidad y las observaciones de Weldon, que nos parecía bastante bueno al principio, empieza a parecernos mucho menos bueno.

(2) Pero ¿quién es responsable de esa desviación sistemática? Los valores teóricos han sido calculados de acuerdo con las reglas del cálculo de probabilidad sobre la base de una cierta asunción, una «hipótesis estadística». No debemos quejarnos de las reglas del cálculo; la falta puede estar en la hipótesis estadística. En efecto, esta hipótesis estadística tiene un punto débil: asumimos que los dados usados en el experimento son «justos». Cuando entre caballeros se juega a los dados se presupone que éstos son «justos», aunque para un naturalista tal asunción no está justificada.

Veamos el ejemplo del físico. Galileo descubrió la ley de caída de los cuerpos que hoy en día escribimos con la notación usual de una ecuación como ésta:

$$s = gt^2/2;$$

donde $s$ denota el espacio (distancia) y $t$ el tiempo. Más exactamente: Galileo descubrió la *forma* de la dependencia de $s$ respecto de $t$: la distancia es proporcional al cuadrado del tiempo $t$. Pero Galileo no hizo ninguna predicción teórica sobre la constante $g$ que entra en esta proporción; el valor adecuado de $g$ debe ser encontrado mediante experimentos. En este aspecto, como en muchos otros, la ciencia natural siguió el ejemplo de Galileo; en innumerables casos la teoría tomó la forma de una ley natural, y el experimento determinaba los valores numéricos de las constantes que entran en la expresión matemática de la ley. Y este procedimiento funciona igualmente en nuestro ejemplo.

Si el dado es «justo», ninguna de sus seis caras es preferible a las

otras, y por ello la probabilidad de que salgan cinco o seis puntos es

$$\frac{2}{6} = \frac{1}{3}.$$

Del mismo modo, aunque el dado no sea «justo» hay cierta probabilidad $p$ para que salgan cinco o seis puntos; $p$ puede ser diferente de 1/3. (Pero no muy diferente en un dado ordinario; de otro modo consideraríamos el dado como «cargado».) Tomemos $p$ como una constante que ha de ser determinada por experimento. Y ahora debemos modificar nuestra hipótesis estadística original: *asumir* que todos (doce) los dados usados tienen la *misma probabilidad p* para exhibir cinco o seis puntos. (Esta es una asunción sencilla, aunque bastante arbitraria. No podemos creer que sea cierta exactamente; sólo podemos esperar que no sea muy diferente de la verdad. No hay probabilidad de que los dados sean exactamente iguales, pero pueden ser ligeramente diferentes.) La otra parte de nuestra hipótesis estadística la mantenemos invariable (dados diferentes y diferentes intentos son considerados como independientes).

Sobre la base de esta nueva hipótesis estadística podemos asignar de nuevo valores teóricos que correspondan a las observaciones apuntadas en la columna (4) de la tabla I. Por ejemplo, el valor teórico que corresponde al valor observado 6.114 es

$$495\, p^4\, (1-p)^8\, 26.306;$$

éste depende de $p$, y lo mismo ocurre con los valores teóricos que corresponden a los otros números de la columna (4).

Queda aún determinar $p$ a partir de los experimentos que estamos examinando. La determinación de $p$ por experimentos no podrá ser exacta, sólo esperamos una aproximación razonable. Si cambiamos nuestro punto de vista por un momento y consideramos la suerte de un solo dado como un ensayo,

$$12 \times 26.306 = 315.672$$

intentos han sido realizados; éste es un número muy grande. La frecuencia del suceso «cinco o seis puntos» puede ser derivada fácilmente de la columna (4) de la tabla I. Encontramos como valor de la frecuencia relativa

$$\frac{106.602}{315.672} = 0.3376986;$$

tomemos esta frecuencia relativa, que resulta de un gran número de intentos, para el valor de $p$. (Asumimos así para $p$ un valor ligeramente más alto que 1/3.)

Una vez $p$ es elegido podemos calcular los valores teóricos que corresponden a las frecuencias observadas. Estos valores teóricos son tabulados en la columna (5) de la tabla I. Así, las columnas (3) y (5) dan los valores teóricos que corresponden a los mismos números observados, aunque calculados bajo hipótesis estadísticas diferentes. En efecto, las dos hipótesis estadísticas difieren sólo en el valor de $p$; la columna (3) utiliza $p = 1/3$, pero la columna (5) utiliza valores ligeramente más altos derivados de las observaciones. (La columna (3) puede ser calculada antes que las observaciones, pero no así la columna (5). Las diferencias entre los datos correspondientes a las columnas (5) y (4) son apuntadas en la columna (6).

No hay duda que los valores teóricos de la columna (5) se adecuan a las observaciones mucho mejor que los de la columna (3). En valores absolutos las diferencias de la columna (6) son, con una sola excepción, menores o iguales a las diferencias de la columna (2) (iguales sólo en tres casos, mucho menores en la mayoría de ellos). En oposición a la columna (2), los signos + y — están entremezclados en la columna (6), de modo que no dan pie a pensar en una desviación sistemática de los valores teóricos de la columna (5) respecto de los datos experimentales de la columna (4).

(3) Juzgada por el ejemplo anterior, la teoría de la probabilidad parece ser bastante adecuada para describir los fenómenos de masas generados por juegos de azar como el de los dados. Si no fuese apropiada para ninguna otra cosa, no merecería mucha atención. Por tanto, consideremos un ejemplo más.

Como informa el Servicio Estadístico oficial de Suiza, hubo trescientos partos de trillizos en dicho país en los treinta años que van de 1871 a 1900. (Es decir, que nacieron novecientos trillizos. Al hablar de partos contamos las madres, no los niños.) El número total de partos (algunos de trillizos, otros de gemelos, la mayoría, por supuesto, de un solo niño) durante el mismo período en la misma unidad geográfica fue de 2.612.246. Así, tenemos aquí un fenómeno de masas de considerables proporciones, pero el suceso considerado, el nacimiento de trillizos, es un *suceso raro*. El número promedio de partos por año es

$$2.612.246/30 = 87.075,$$

el número promedio de partos de trillizos sólo

$$300/30 = 10.$$

Desde luego, el suceso ocurrió más a menudo en ciertos años, en otros menos a menudo que el promedio 10 y en algunos años diez veces exactamente. La tabla II da los detalles relevantes en la columna (2). Vemos aquí (en la fila que tiene 10 en la primera columna) que hubo, en el período considerado, cuatro años en que se dieron diez partos de trillizos. Como muestra la misma columna (2) ningún año del período tuvo menos de tres partos tales, ninguno tuvo más de diecisiete y cada uno de estos números extremos, 3 y 17, se dieron en sólo un año.

Los números de la columna (2) parecen estar dispersos de alguna manera azarosa. Es interesante notar que el cálculo de probabilidad es capaz de igualar los números observados que parecen irregulares en la columna (2) por medio de números teóricos que siguen una sencilla ley; ver la columna (3). El acuerdo entre las columnas (2) y (3), juzgado por inspección, no parece ser malo; la diferencia entre los dos números, los observados y los teóricos, es menos de uno en valor absoluto, excep-

TABLA II

TRILLIZOS NACIDOS EN SUIZA

(1871-1900)

| (1) Partos | (2) Años observados | (3) Años teóricos | (4) (2) acumulado | (5) (3) acumulado |
|---|---|---|---|---|
| 0 | 0 | 0,00 | 0 | 0,00 |
| 1 | 0 | 0,00 | 0 | 0,00 |
| 2 | 0 | 0,09 | 0 | 0,09 |
| 3 | 1 | 0,21 | 1 | 0,30 |
| 4 | 0 | 0,57 | 1 | 0,87 |
| 5 | 1 | 1,14 | 2 | 2,01 |
| 6 | 1 | 1,89 | 3 | 3,90 |
| 7 | 5 | 2,70 | 8 | 6,60 |
| 8 | 1 | 3,39 | 9 | 9,99 |
| 9 | 4 | 3,75 | 13 | 13,74 |
| 10 | 4 | 3,75 | 17 | 17,49 |
| 11 | 4 | 3,42 | 21 | 20,91 |
| 12 | 3 | 2,85 | 24 | 23,76 |
| 13 | 2 | 2,16 | 26 | 25,92 |
| 14 | 1 | 1,59 | 27 | 27,51 |
| 15 | 2 | 1,02 | 29 | 28,53 |
| 16 | 0 | 0,66 | 29 | 29,19 |
| 17 | 1 | 0,39 | 30 | 29,58 |
| 18 | 0 | 0,21 | 30 | 29,79 |
| 19 | 0 | 0,12 | 30 | 29,91 |

to en dos casos. Pero en estos dos casos (las filas con siete y ocho en la primera columna) la diferencia es mayor de dos en valor obsoluto.

Hay una estratagema que nos permite juzgar un poco mejor el acuerdo de las dos series de números. La columna (4) de la tabla II contiene los números de la columna (2) «cumulativamente». Por ejemplo, consideremos la fila que tiene siete en la columna (1); ésta tiene cinco en la columna (2) y ocho en la columna (4). Ahora bien,

$$8 = 0 + 0 + 0 + 1 + 0 + 1 + 1 + 5;$$

es decir, 8 es la suma o la «acumulación» de todos los números de la columna (2) hasta el número 5 inclusive, en la fila respectiva. (En otras palabras: ocho es el número de los años del período en que el número de partos de trillizos *no excedió* a siete.) La columna (5) contiene los números de la columna (3) «cumulativamente», y así, las columnas (4) y (5) son análogamente derivadas de los números observados de la columna (2) y los números teóricos de la columna (3), respectivamente. El acuerdo entre las columnas (4) y (5) parece excelente; la diferencia es menos de uno, excepto en un solo caso en que es todavía menor de dos.

## 6. Explicación de fenómenos

Las ideas conectadas con el concepto de probabilidad juegan un papel en la explicación de fenómenos, y esto es cierto con relación a fenómenos de cualquier ciencia, desde la física a las ciencias sociales. Consideremos dos ejemplos.

(1) Gregor Mendel (1822-1884), experimentando con el cruce de plantas, encontró una nueva ciencia, la genética. Mendel fue, por cierto, abad en Moravia, y realizó sus experimentos en el jardín de su monasterio. Su descubrimiento, aunque muy importante, es sencillo de veras. Para comprenderlo nos basta con la descripción de un experimento y una noción intuitiva de la probabilidad. Para hacer las cosas aún más fáciles no examinaremos ninguno de los propios descubrimientos de Mendel, sino otro realizado por uno de sus seguidores *.

De dos plantas íntimamente relacionadas (especies diferentes del mismo género) una tiene flores blancas y la otra flores rojo oscuro. Las dos plantas son tan próximas que la una puede fertilizar a la otra. Las semillas que resultan de tal cruce dan plantas híbridas de carácter

* Por Correns; ver W. Johannsen, *Elemente der exakten Erblichkeitslehre*, Jena, 1909, pág. 371.

intermedio: los híbridos tienen flores rosas. (En la figura 14.3 el rojo es indicado por un sombreado más intenso; el rosa, por otro menos intenso.) Si se permite fertilizarse a estas plantas híbridas las semillas que resultan dan a la tercera generación plantas de las tres clases: plantas con flores blancas, rosas y rojas. La figura 14.3 representa esquemáticamente las relaciones entre las tres generaciones subsecuentes.

Fig. 14.3. Tres generaciones en un experimento mendeliano.

Pero el rasgo más chocante del fenómeno es la proporción numérica en que se producen las tres clases de plantas de la tercera generación. En el experimento descrito, 564 plantas de la tercera generación se han observado. Entre éstas, las dos clases de plantas que se parecen a uno u otro de los abuelos son igualmente numerosas: hay 141 plantas con flores blancas y 132 con flores rojas en la tercera generación. Pero las plantas que se parecen a los padres híbridos resultan mucho más numerosas: hay 291 plantas con flores rosas en la mencionada tercera generación. Podemos examinar estos números en la figura 14.3. Por nuestra parte, notemos que estos números dados por el experimento están en una proporción sencilla:

141 : 291 : 132 casi como 1 : 2 : 1.

Esta sencilla proporción invita a una explicación sencilla.

Empecemos por el principio. El experimento empezó con el cruce de dos clases de plantas. Todas las plantas con flores surgen de la unión de dos células reproductivas (un óvulo y un grano de polen). Los híbridos de flor rosa de la segunda generación surgen de dos células reproductivas de extracción diferente. Puesto que las plantas de flor rosa de la tercera generación son semejantes a las de la segunda generación, es natural asumir que se han engendrado de modo semejante, por dos células reproductivas de clase *diferente*. Esto nos conduce a suponer que los híbridos de flor rosa de la segunda generación *tienen* dos clases diferentes de células reproductivas. No obstante, al suponer esto percibimos una posibilidad de explicar el cruce salteado. En efecto, veríamos más claramente lo que sucede si los híbridos de flor rosa de la segunda generación tuviesen, en realidad, dos clases diferentes de células reproductivas, a las que podemos llamar células «roja» y «blanca». Cuando dos células tales se combinan, la combinación puede ser blanco con blanco, rojo con rojo o un color con otro, y estas tres combinaciones diferentes *podrían* explicar las tres clases diferentes de plantas en la tercera generación; ver la figura 14. 3.

Tras esta observación no sería difícil explicar las proporciones numéricas. La desviación de la proporción observada 141 : 291 : 132 respecto de la proporción 1 : 2 : 1 aparece como azarosa; es decir, parece semejante a la desviación de las frecuencias observadas respecto de las probabilidades subyacentes. Esto nos lleva a preguntarnos lo que son las probabilidades de dos clases de células, o en qué proporción se engendran las células «rojas» y «blancas». Como hay, aproximadamente, tantas plantas de flor blanca como de flor roja en la tercera generación, podemos guardarnos de intentar lo más sencillo: asumimos que las plantas de flor rosa engendran células reproductivas «blancas» y «rojas» en igual número. Finalmente, nos vemos casi conducidos a comparar el azaroso encuentro de dos células reproductivas con la extracción azarosa de dos bolas, y así llegamos al siguiente y sencillo problema.

Hay dos sacos que contienen bolas blancas y rojas, sin ninguna más de otro color. Cada saco contiene justo tantas bolas blancas como rojas. Con ambas manos extraigo de los dos sacos una bola de cada. Encontrar la probabilidad de sacar dos bolas blancas, dos bolas de colores diferentes y dos rojas.

Como puede verse fácilmente (cf. la sección 3 (5)), las probabilidades requeridas son

$$\frac{1}{4}, \frac{2}{4}, \frac{1}{4},$$

respectivamente. Ahora percibimos una sencilla razón para la proporción 1 : 2 : 1 que parece subyacer a los números observados, y al hacerlo así nos aproximamos mucho a los conceptos esenciales de Mendel.

(2) El concepto de los fenómenos azarosos de masas juega un papel importante en la física. Al objeto de ilustrar este papel consideremos la velocidad de las reacciones químicas.

Observaciones relativamente toscas son suficientes para sugerir que la velocidad de un cambio químico depende de la concentración de las sustancias reactivas. (Por concentración de una sustancia entendemos cantidad por unidad de volumen.) Esta dependencia de la velocidad de reacción química respecto de la concentración de los reactivos fue reconocida muy pronto, pero el descubrimiento de la forma matemática de la dependencia vino mucho después. Un caso particular importante fue observado por Wilhelmy en 1850, mientras que la ley general fue descubierta por dos químicos noruegos, Guldberg y Waage, en 1867. Ahora nosotros bosquejaremos, en un caso particular y del modo más sencillo posible, algunas de las consideraciones que condujeron a Guldberg y Waage a su descubrimiento.

Consideremos una reacción bimolecular. Es decir, dos sustancias diferentes, $A$ y $B$, participan en la reacción que consiste en la combinación de una molécula de la primera sustancia $A$ con una molécula de la segunda sustancia $B$. Las sustancias $A$ y $B$ son disueltas en agua, y el cambio químico ocurre en esta solución. Las sustancias que resultan de la reacción no participan en la acción química; éstas son inactivas en uno y otro sentido. Por ejemplo, pueden ser insolubles en agua y depositadas en forma sólida.

La solución en que ocurre la reacción consta de un gran número de moléculas. De acuerdo con las ideas de los físicos (la teoría cinética de la materia), estas moléculas están en movimiento, viajando a varias velocidades, algunas muy grandes, y chocando una y otra vez. Si una molécula $A$ choca con una molécula $B$ pueden verse de tal modo envueltas que cambien algunos de sus átomos: la reacción química de que tratamos aquí consiste en un cambio tal, suponemos. Quizá es necesario para tal cambio que las moléculas choquen a velocidad muy alta o que estén dispuestas en una posición favorable de la una con respecto a la otra en el momento de su colisión. De cualquier modo, cuanto más a menudo choca una molécula $A$ con una molécula $B$ más oportunidad hay para la combinación química de dos moléculas tales, y más alta será la velocidad de la reacción química. Y así nos vemos conducidos a la siguiente conjetura: ***la velocidad de reacción es proporcional al número de colisiones entre las moléculas A y las moléculas B.***

Nos resulta imposible predecir exactamente el número de tales colisiones. Tenemos ante nosotros un fenómeno azaroso de masas como el del chaparrón. Recordemos la figura 14.2; tampoco entonces podíamos predecir exactamente cuántas gotas de lluvia caen en el subárea $F$. Sin embargo, podíamos predecir que el número de gotas de lluvia que chocan en el subárea $F$ sería *proporcional* al número de gotas que caen en el área total $P$. (La proporcionalidad es aproximada, y el factor de proporcionalidad es $F/P$, como discutimos al final de la sección 2.) De modo semejante podemos predecir que el número de colisiones en que estamos interesados (entre cualesquiera moléculas $A$ y $B$) será proporcional al número de las moléculas $A$. Por supuesto, será también proporcional al número de moléculas $B$, y, finalmente, proporcional al *producto* de estos dos números. Pero el número de moléculas de una sustancia es proporcional a la concentración de dicha sustancia, y así, nuestra conjetura nos conduce a este enunciado: la *velocidad de reacción es proporcional al producto de las concentraciones.*

Hemos llegado a un caso particular de la ley general de la acción química de masas descubierta por Guldberg y Waage. Este es el caso particular apropiado para las consideradas circunstancias particulares. Sobre la base de la ley de acción de masas es posible calcular la concentración de las sustancias reactivas en cualquier momento y predecir el curso total de la reacción.

## 7. Juzgando hipótesis estadísticas

Empecemos con una anécdota [*].

(1) «Un día en Nápoles el reverendo Galiani vio desde la basílica a un hombre que, moviendo tres dados en un cubilete, apostó que sacaba tres seises; y, en efecto, los sacó. Tanta suerte es imposible, dirá usted. Pero el hombre lo consiguió una segunda vez, y apostó que lo repetiría. Metió los dados en el cubilete, tres, cuatro, cinco veces y cada vez sacó tres seises. «¡Sangre de Baco! —exclamo el reverendo—, ¡los dados están cargados!» Y, efectivamente, lo estaban. Pero ¿por qué usó el reverendo una expresión violenta?».

El reverendo Galiani sacó una conclusión plausible de un tipo muy importante. Si descubrió por sí mismo este importante tipo de inferencia plausible por el estímulo de la situación, su excitación es bastante comprensible, y yo, personalmente, no le reprocharía la violencia de su expresión.

---

[1] J. Bertrand, *Calcul des probabilités*, págs. VII-VIII.

Lo correcto es tratar a todo el mundo como un caballero mientras no se tenga alguna evidencia definitiva de lo contrario. De modo semejante, lo correcto es empezar un juego de azar bajo la suposición de que es un juego limpio. No dudo de que el reverendo hizo lo correcto y asumió al principio que el hombre que él vio desde la basílica tenía dados «justos» y los usaba «limpiamente». Tal asunción, correctamente establecida en términos de probabilidad, es una hipótesis estadística. Una hipótesis estadística generalmente asume los valores de ciertas probabilidades. Así, el reverendo asumió, más o menos explícitamente, que cualquiera de los dados envueltos mostrará seis puntos con la probabilidad de 1/6. (Tenemos aquí la misma hipótesis estadística que en la sección 5 (1).)

El cálculo de probabilidad nos capacita para calcular probabilidades deseadas a partir de probabilidades dadas, sobre la base de una hipótesis estadística dada. Así, pues, sobre la base de la hipótesis estadística adoptada al principio por el reverendo podemos calcular la probabilidad de que salgan tres seises con tres dados; ésta es

$$(1/6)^3 = 1/216,$$

una probabilidad bastante pequeña. La probabilidad de repetir esta suerte dos veces —es decir, sacar tres seises en un primer intento, y sacarlos de nuevo en un segundo intento— es

$$(1/216)^2 = (1/6)^6 = 1/46.656,$$

una probabilidad muy pequeña realmente. Pero el hombre de marras repitió el mismo rasgo extraordinario cinco veces. Enumeremos las probabilidades correspondientes:

| *Repeticiones* | *Probabilidad* | | |
|---|---|---|---|
| 1 | $1/6^3$ | = | 1/216 |
| 2 | $1/6^6$ | = | 1/46,656 |
| 3 | $1/6^9$ | = | 1./10,077,696 |
| 4 | $1/6^{12}$ | = | 1/2,176,782,336 |
| 5 | $1/6^{15}$ | = | 1/470,184,984,576. |

Quizá el reverendo adoptó su suposición inicial por mera cortesía; pues viendo al hombre desde la escalinata podía ya sospechar de la honradez de los dados. El reverendo quedó silencioso cuando los tres seises salieron en doble sucesión, un hecho que bajo la creencia inicial no ocurriría mucho más que una vez en 50.000 intentos. El permaneció

silencioso aún más. Pero, cuando el suceso se fue haciendo más y más improbable, alcanzó y sobrepasó el grado de probabilidad que la gente considera como milagroso el reverendo perdió la paciencia, sacó su conclusión, rechazó su cortés suposición inicial y soltó su lengua.

(2) La anécdota que hemos examinado arriba es interesante en un solo aspecto: es típica. Muestra claramente las circunstancias bajo las que puede rechazarse razonablemente una hipótesis estadística. Saquemos consecuencias de la hipótesis estadítica propuesta. De interés especial son las consecuencias concernientes con algún suceso que parece *muy improbable* desde el punto de vista de nuestra hipótesis estadística; quiero decir, un suceso cuya probabilidad, calculada sobre la base de hispótesis estadísticas, es muy pequeña. Ahora apelemos a la experiencia: observemos un intento que pueda producir ese suceso alegadamente improbable. Si el suceso, a despecho de su baja probabilidad calculada, sucede, en realidad, constituirá un fuerte *argumento contra* la hipótesis estadística propuesta. En efecto, encontramos dificil creer que nada tan extremadamente improbable pueda suceder. Pero, innegablemente, la cosa sucedió. Entonces nos damos cuenta de que cualquier probabilidad es calculada sobre la base de una hipótesis estadística y empezamos dudando de las bases del cálculo de esa pequeña probabilidad. Y así surge el argumento contra la subyacente hipótesis estadística.

(3) Como el reverendo Galiani, nos sentimos obligados a rechazar la hipótesis de los dados «justos» cuando observamos las extensas observaciones referidas en la sección 5 (1); no obstante, nuestras razones no son tan agudas como las suyas. ¿Podríamos encontrar mejores razones a la luz de la discusión anterior?

Aquí están los hechos: 315.672 intentos de que salgan cinco o seis puntos con un dado logran 106.602 éxitos; ver la sección 5 (2). *Si* el dado fuese «justo», la probabilidad de éxito sería 1/3. Por tanto, esperaríamos

$$315.672/3 = 105.224$$

éxitos en 315.672 intentos. Así, el número observado se desvía del número esperado

$$106.602 - 105.224 = 1.378$$

unidades. ¿Tal hipótesis habla a favor o en contra de la hipótesis del dado «justo»? ¿Consideraremos la desviación 1.378 como grande o pequeña? ¿Es la probabilidad de tal desviación alta o baja?

La última pregunta parece ser la sensata. Sin embargo, aún necesitamos una interpretación razonable de la breve, aunque importante,

palabra «tal». Rechazaremos la hipótesis estadística si la probabilidad que estamos por calcular resulta ser baja. Sin embargo, la probabilidad de que la desviación sea igual a 1.378 unidades es muy pequeña de cualquier modo —incluso la probabilidad de una desviación exactamente igual a 0 sería muy pequeña—. Por tanto, tenemos que tener en cuenta *todas las desviaciones del mismo, o mayor valor absoluto que la desviación observada.* Y así, nuestro juicio depende de la solución del siguiente problema: *Dado que la probabilidad de éxito es 1/3 y que los intentos son independientes, encontrar la probabilidad de que en 315.672 intentos el número de éxitos sea o mayor que 106.601 o menor que 103.847.*

Con algún conocimiento del cálculo de probabilidad encontramos que la probabilidad requerida es, aproximadamente,

$$0.0000001983;$$

esto significa menos de dos oportunidades en un millón. Es decir, un suceso que parece extremadamente importante ha ocurrido *si* se acepta que la hipótesis estadística subyace al cálculo de probabilidad. Encontramos difícil de creer que un suceso tan improbable haya ocurrido en realidad, y así, la hipótesis subyacente del dado «justo» parece extremadamente inverosímil. En la sección 5 (1) ya vimos una buena razón para rechazar la hipótesis del dado «justo», pero ahora vemos una todavía mejor, más clara.

(4) El acontecimiento real de un suceso al que una cierta hipótesis estadística atribuye una pequeña probabilidad es un argumento contra la hipótesis, y cuanto más pequeña la probabilidad, más fuerte el argumento.

Al objeto de visualizar este punto esencial consideremos la secuencia

$$\frac{1}{10}, \frac{1}{100}, \frac{1}{1.000}, \frac{1}{10.000}, \cdots$$

Una hipótesis estadística implica que la probabilidad de un cierto suceso es 1/10. El suceso ocurre. ¿Rechazaremos la hipótesis? Bajo circunstancias normales, la mayoría de nosotros no nos sentiríamos autorizados para rechazarlo; el argumento contra la hipótesis no nos parece bastante fuerte. Si la probabilidad alegada es 1/1.000, y aun el suceso ocurre, el caso es todavía más fuerte contra la hipótesis. Si la hipótesis estadística atribuye la probabilidad

$$\frac{1}{1.000.000.000}$$

al suceso, o una oportunidad en un billón, a pesar de lo cual el suceso ocurre, casi todo el mundo consideraría la hipótesis como desacreditada sin esperanza, aunque no haya necesidad lógica de rechazar la hipótesis en este punto. Si, no obstante, la secuencia procede sin interrupción de modo que el suceso ocurre una vez tras otra de modo que la hipótesis estadística atribuye probabilidades continuamente decrecientes hacia 0, toda persona razonable llega a un momento crítico en que se siente justificada a rechazar la hipótesis, haciéndose insostenible por el aumento de sus consecuencias improbables. Y este punto es claramente sugerido por la historia del reverendo Galiani. La probabilidad de una primera salida de tres seises era 1/216; de la secuencia de cinco salidas de tres seises cada una, 1/470.184.984.576.

La precedente discusión es de especial importancia para nosotros si adoptamos el punto de vista de que la teoría de la probabilidad es una parte del estudio de la naturaleza. Toda ciencia natural debe recurrir a las observaciones. Por tanto, debe adoptar reglas que especifiquen de algún modo las circunstancias bajo las que sus enunciados son confirmados o refutados por la experiencia. Hemos hecho esto para la teoría de la probabilidad. Hemos descrito ciertas circunstancias bajo las que podemos considerar razonablemente una hipótesis estadística como prácticamente refutada por las observaciones. Por otro lado, si una hipótesis sobrevive a varios intentos de refutación, podemos considerarla corroborada hasta cierto punto.

(5) La probabilidad, según ha sido definida en la sección 2, es el valor teórico de una frecuencia relativa de largo alcance. Lo anterior nos dio la oportunidad de comprobar unas cuantas cosas. Primero, un valor teórico tal depende, por supuesto, en nuestra teoría, de nuestras asunciones iniciales, la hipótesis estadística adoptada. Segundo, tal valor teórico puede ser muy diferente del valor real.

Una conveniente notación puede ayudarnos a aclarar nuestras ideas. Sea $P$ la probabilidad de un suceso $S$ calculado sobre la base de una cierta hipótesis estadística $H$. Luego $P$ depende tanto de $S$ como de $H$. (De hecho, podríamos usar, en vez de $P$, el símbolo más explícito $P\,(S, H)$ que enfatiza la dependencia de $P$ respecto de $S$ y $H$.)

En algunas de las anteriores aplicaciones tomamos la hipótesis $H$ como garantizada (al menos por el momento) y, calculando $P$ sobre la base de $H$, intentamos predecir la frecuencia observable del suceso $S$. Sin embargo, en la presente sección hemos procedido en otra dirección. Habiendo observado el suceso $S$, calculamos $P$ sobre la base de la hipótesis estadística $H$ y, a la vista del valor $P$ obtenido, intentamos juzgar el grado de confianza de la hipótesis $H$. Cuanto *más pequeña* es $P$,

más inclinados nos sentimos a rechazar la hipótesis $H$ y *más inverosímil* se nos muestra ésta: $P$ indica la posibilidad de la hipótesis $H$. De aquí en adelante diremos que $P$ es la *verosimilitud de la hipótesis estadística* $H$, juzgada en vista de que $S$ ha sido observado.

Esta terminología, que concuerda esencialmente con los usos de los estadísticos, hace hincapié en cierto aspecto de la dependencia de $P$ respecto del suceso $S$ y de la hipótesis estadística $H$. Nuestra terminología original pone el acento sobre el aspecto complementario de la misma dependencia: $P$ es la *probabilidad del suceso* $S$ calculado sobre la base de la hipótesis estadística $H$.

Se necesita alguna práctica en el uso de esta doble terminología para convencernos de que sus ventajas superan suficientemente sus peligros.

## 8. Eligiendo entre hipótesis estadísticas

El ejemplo siguiente nos da una primera orientación para las aplicaciones de la teoría de la probabilidad en la investigación estadística.

(1) Un consumidor compra de un productor un cierto artículo al por mayor. El consumidor es importante: un gran traficante o firma industrial, o una agencia del Gobierno. El productor es también importante y manufactura el artículo en cuestión a gran escala. El artículo puede ser clavos, o pestillos, o cualquier otra manufactura; un ejemplo interesante es una espoleta, usada por explosivos de fuego en operaciones de carga o aprovisionamiento. El artículo debe reunir ciertas especificaciones. Por ejemplo, los clavos no deben ser más largos de 2,04 pulgadas ni más cortos de 1,96, su espesor debe ser especificado del mismo modo y quizá también la fuerza mínima necesaria para la penetración del clavo; el tiempo de combustión de la espoleta es especificado también, etc. Un artículo que no reúne las especificaciones se considera como *defectuoso*. Aun la remesa más cuidadosamente manufacturada puede contener una pequeña fracción de defectuosos. Por tanto, la remesa debe ser inspeccionada antes de que pase del productor al consumidor. Puede ser totalmente inspeccionada, esto es, hay que probar si cada artículo reúne las especificaciones acordadas. Tal inspección sería impracticable para una remesa de 10.000 clavos y sería absurda en el caso de una remesa de espoletas aun cuando ésta fuera pequeña, pues para medir el tiempo de combustión habría que destruir la espoleta y no es cuestión de destruir toda la remesa para inspeccionarla. Por tanto, en muchos casos, en vez de inspeccionar toda la serie antes de su aceptación se toma una muestra relativamente pequeña. Un procedimiento sencillo de muestreo se caracteriza por la siguiente regla.

«Tomemos al azar una muestra de $n$ artículos de la serie de $N$ artículos que tenemos a la vista. Probemos cada artículo de la muestra. Si el número de piezas defectuosas de la muestra no excede un cierto número acordado $c$, el llamado *número de aceptación*, el consumidor acepta la compra, pero lo rechaza, y el vendedor tendrá que llevárselo, si hay más artículos defectuosos de $c$ en la muestra.»

Los resultados obtenidos con esta regla dependen de la suerte, pues la fracción de artículos defectuosos de la muestra puede ser mucho más alta o mucho más baja que en el conjunto total. Si la muestra resulta ser mejor que el total la suerte está en contra del consumidor y está en contra del productor si la muestra resulta peor que el total. A despecho de estos riesgos, tal procedimiento parece necesario, y la regla formulada puede ser bastante razonable. Tenemos que descubrir cómo funciona el procedimiento, cómo su resultado depende de la calidad de la mercancía de que se trata. Y así somos llevados a formular el siguiente problema: *Dada la probabilidad $P$ de que un artículo elegido al azar en una remesa sea defectuoso, encontrar la probabilidad $a$ de que tal remesa sea aceptada.*

En los más importantes casos prácticos $N$, el tamaño de la remesa, es grande incluso en comparación con $n$, el tamaño de la muestra. En tales casos asumimos que $N$ es infinito; con ello perdemos poco en precisión y ganamos mucho en simplicidad. Asumiendo que $N = \infty$, encontramos fácilmente que

$$a = (1-p)^n + \binom{n}{1} p(1-p)^{n-1} + \binom{n}{2} p^2 (1-p)^{n-2} + \ldots + \binom{n}{c} p^c (1-p)^{n+c}.$$

Tomemos esta expresión de la probabilidad de aceptación $a$ como garantizada y concentrémonos en la discusión de alguna de sus implicaciones prácticas.

Hacemos el diagrama de $a$ en función de $p$; ver la figura 14.4. Si hacemos el diagrama de 100 $a$ como función de 100 $p$, la forma de la curva será la misma. Ahora bien, 100 $p$ es el porcentaje de artículos defectuosos en la remesa suministrada. Por otro lado, remesas con el mismo porcentaje de objetos defectuosos están sometidas al mismo procedimiento de inspección; la frecuencia relativa de aceptación, es decir, la razón de remesas aceptada con respecto a remesas suministradas, será próxima a $a$. Por tanto, a largo plazo, 100 $a$ será el porcentaje de remesas aceptadas entre las mismas suministradas. Esto explica las etiquetas

puestas en los ejes de la figura 14.4. La curva de la figura 14.4 nos permite examinar cómo opera el procedimiento en remesas de artículos cualitativamente distintos, y así resulta llamada apropiadamente *característica de operaciones.*

Juzgado por sus efectos, ¿parece razonable el procedimiento? Esta es la cuestión que deseamos considerar.

Si no hay piezas defectuosas en la serie, no hay posibilidad de rechazarla. En efecto, si $p = 0$ nuestra fórmula resulta $a = 1$. Si sólo hay piezas defectuosas en la serie, no habrá posibilidad de aceptarla. De hecho, si $p = 1$ nuestra fórmula será $a = 0$. Ambos casos extremos de la curva son obvios.

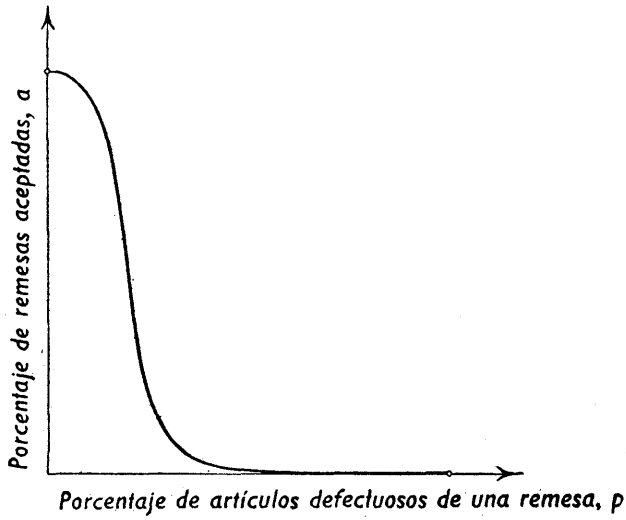



Fig. 14.4. Característica de operaciones en un procedimiento de muestreo para la aceptación de mercancías.

Si el número de artículos defectuosos aumenta, las posibilidades de aceptación disminuirán. Diferenciando con un poco de habilidad encontramos fácilmente la sorprendentemente sencilla expresión

$$\frac{da}{dp} = -(n-c)\binom{n}{c}p^c(1-p)^{n-1-c},$$

que es siempre negativa. Por tanto, la característica de operaciones es necesariamente una curva descendente, como la representada en la figura 14.4, que es como debe ser.

El valor absoluto de la derivada, o $-da/dp$, tiene también una

cierta significación práctica. El cambio $dp$ de la abscisa representa un cambio en la calidad de la mercancía. El cambio $da$ de la ordenada representa un cambio en las posibilidades de aceptación, debido al cambio de calidad. Cuanto mayor es la razón de estas posibilidades $da/dp$ en valor absoluto más aguda es la distinción hecha por el procedimiento entre dos remesas ligeramente diferentes. Especialmente, el punto en que $da/dp$ alcanza su máximo valor absoluto puede ser llamado con propiedad el «punto de más aguda discriminación». Este punto es fácilmente recognoscible en el diagrama: se trata del punto de inflexión, si es que le hay, y si no, el extremo izquierdo de la curva. (Su abscisa es $p = c/(n - 1)$.

(2) La regla se muestra sensata también desde otro punto de vista. Tiene cierta flexibilidad. Eligiendo $n$, el tamaño de la muestra, y $c$, el número de aceptación, podemos adaptar la regla de las demandas concretas. Tanto el consumidor como el productor requieren protección contra los riesgos inherentes al muestreo. Una mala remesa puede a veces dar una buena muestra y una buena remesa puede dar una mala muestra, y así hay dos clases de riesgos: el procedimiento de muestreo puede hacer que se acepte una mala mercancía o que se rechace una buena. El consumidor está en contra de aceptar mala mercancía y el productor en contra de que se rechace la buena. Las dos cosas poco deseables pueden ocurrir y lo único que podemos pedir es que no ocurran demasiado a menudo. Esto nos lleva a problemas concretos como el siguiente.

«Determinar el tamaño de la muestra y el número de aceptación de modo que haya menos de una oportunidad entre diez de que una remesa con el 5 por 100 de artículos defectuosos sea aceptada y que haya menos de cinco oportunidades entre cien de que una remesa con sólo el 2 por 100 de artículos defectuosos sea rechazada.»

En este problema hay dos incógnitas, el tamaño de la muestra $n$ y el número de aceptación $c$. La condición del problema requiere las dos desigualdades siguientes:

$$a > 0{,}95 \quad \text{cuando } p = 0{,}02,$$

$$a < 0{,}1 \quad \text{cuando } p = 0{,}05.$$

Es posible satisfacer estas dos condiciones simultáneas, pero hay que hacer muchos números para encontrar el tamaño más bajo de la muestra $n$ y el número de aceptación correspondiente $c$ para que las desigualdades requeridas se mantengan.

No examinaremos la parte numérica de la cuestión. Estamos más interesados en visualizar el problema que en resolverlo. Por tanto,

miremos un poco más su base. Como ya hemos dicho, tanto la aceptación de una mala remesa como el rechazo de otra buena son igualmente indeseables, lo primero desde el punto de vista del consumidor, lo segundo desde el del productor. Sin embargo, aunque ambas posibilidades sean indeseables pueden no serlo igualmente y los intereses del consumidor y del productor pueden no estar tan agudamente opuestos. La aceptación de un mal lote puede estar en contra del interés del productor y dañar así su reputación. Pero el rechazo de uno bueno estará mucho más en contra de los intereses del consumidor; puede ser que necesite los artículos urgentemente y el rechazo retrasar el pedido de modo considerable. Además, repetidos rechazos de buenos lotes, o el simple peligro de tal rechazo, puede subir el precio. Si los intereses de ambas partes están de acuerdo, el rechazo de un buen lote puede ser menos deseable que la aceptación de uno malo. Una vez visto su fundamento parece comprensible que las condiciones de nuestro problema procuran más protección contra el rechazo de buenas mercancías que contra la aceptación de las malas. (Sólo cinco oportunidades entre cien se permiten para el primer suceso indeseable, mientras para el segundo se admiten diez entre cien.)

(3) El problema examinado en (2) admite otra interpretación de algún modo diferente.

El abogado del productor afirma que no hay más del 2 por 100 de artículos defectuosos en la remesa. Pero el abogado del consumidor mantiene que hay por lo menos el 5 por 100 de artículos defectuosos en ella. Por alguna razón (puede ser un lote de espoletas), una inspección completa no tiene lugar; por tanto, algún procedimiento de muestreo debe decidir entre las dos afirmaciones. Para este propósito, el procedimiento bosquejado en (1) con los datos numéricos de (2) puede ser usado convenientemente.

De hecho, las afirmaciones en conflicto de dos abogados sugieren una ficción. Podemos pretender que hay dos posibilidades con respecto al lote: el porcentaje de objetos defectuosos es exactamente el 2 por 100 o exactamente el 5 por 100. Por supuesto que nadie cree tal ficción, pero el estadístico puede encontrarla conveniente: restringe su tarea a la elección entre dos alternativas claras y sencillas. Si las partes están de acuerdo en que el rechazo de una remesa con el 2 por 100 de objetos defectuosos es menos deseable que la aceptación de otra con el 5 por 100 de objetos defectuosos el estadístico puede esperar razonablemente adoptar el procedimiento delineado en (1) con los datos numéricos prescritos en (2). Si la elección del estadístico satisface a los abogados o a los filósofos, no me aventuro a decirlo, pero, desde luego, tiene una clara relación con los hechos del caso. La regla del estadístico aplicada

a un gran número de casos análogos acepta un buen lote (con 2 por 100 defectuoso) unas novecientas cincuenta veces de mil y lo rechaza sólo unas cincuenta veces, pero la misma regla rechaza un mal lote (con el 5 por 100 defectuoso) unas novecientas de mil y lo acepta sólo unas cien veces. Esto es, la regla del estadístico, basada en el muestreo, no puede esperar dar siempre una recta decisión, pero puede esperar darla en un considerable porcentaje de casos *a largo plazo.*

(4) Intentar dar una idea de lo que hace el estadístico con un solo ejemplo es, naturalmente, una empresa desesperada. Pero el anterior ejemplo puede darnos una cierta idea de la tarea del estadístico que, aunque muy incompleta, no resulta demasiado distorsionada: el estadístico inventa reglas de la misma naturaleza que el procedimiento de aceptación de muestras bosquejando en (1) y considerado en lo referente a datos numéricos en (2). Podemos comprender la tarea del estadístico si comprendemos la naturaleza de las reglas que él descubre. Por tanto, tenemos que formular en términos generales lo que parece ser esencial en nuestra regla particular; quiero decir, la regla examinada en las subsecciones (1), (2) y (3).

Nuestra regla prescribe una elección entre dos cursos de acción, aceptación y rechazo. Sin embargo, el aspecto del problema considerado en (3) es adecuado para la generalización. Consideramos entonces una *elección entre dos hipótesis estadísticas.* (Estas son «una muestra casual tomada de una gran remesa con el 2 por 100 defectuoso» y «una muestra casual tomada de una gran remesa con el 5 por 100 defectuoso».) Toda elección razonable deberá ser hecha teniendo en cuenta la experiencia pasada y las consecuencias futuras. En efecto, nuestra regla es trazada con relación a ambas.

Según nuestra regla, la elección depende de un conjunto de observaciones claramente especificadas (la comprobación de $n$ artículos y el número de los defectuosos detectados entre esos $n$ artículos). Estas observaciones constituyen la experiencia relevante en que está basada la elección. Como nuestra regla prefiere una hipótesis a otra a base de observaciones, bien pudiera ser llamada una regla *inductiva.*

Nuestra regla es trazada con miras a las consecuencias probables. El estadístico no puede predecir las consecuencias de una aplicación individual de la regla. Sólo puede predecir cómo obrará la regla *a largo plazo.* Si la elección prescrita por la regla es ensayada muchas veces en tales y tales circunstancias, nos conducirá a tal y tal resultado, en tal y tal porcentaje de ensayos, a largo plazo. Nuestra regla está trazada con miras a *consecuencias de largo alcance.*

En resumen, nuestra *regla está trazada para elegir entre hipótesis*

*estadísticas, y, basada en un conjunto específico de observaciones, apunta a consecuencias de largo alcance.* Si podemos mirar nuestra regla como suficientemente típica tenemos una idea de lo que los estadísticos están haciendo: trazar reglas de esta clase.

(En efecto, ellos intentan crear las «mejores» reglas de esta clase. Por ejemplo, desean hacer un mínimo de las posibilidades de tal y tal efecto indeseable, siendo dado el tamaño de la muestra, del que dependen el trabajo y los gastos de las observaciones.)

(5) Tomar al azar una muestra de un lote es una operación importante en la investigación estadística. Hay aún otro problema en esta operación, que discutiremos aquí. Conservaremos nuestra notación anterior al formular el problema.

*En un lote muy grande, el 100 por 100 de artículos es defectuoso. Al objeto de obtener alguna información de p, tomemos una muestra de n artículos del lote, entre los que se encuentran m defectuosos. Sobre esta observación, ¿qué valor atribuir razonablemente a p?*

Hay una contestación obvia, sugerida por la misma definición de la probabilidad. Sin embargo, el problema es importante y merece ser examinado desde varios ángulos.

Nuestra observación nos da alguna información de $p$. Especialmente, si $m$ resulta ser diferente de 0, concluiremos que $p$ es diferente de 0. De modo semejante, si $m$ es menos de $n$, concluiremos que $p$ es menos de 1. Sin embargo, en cualquier caso $p$ continua siendo una incógnita y todos los valores entre 0 y 1 son elegibles para $p$. Si atribuimos uno de estos valores a $p$, hacemos una conjetura por intuición, elegimos una hipótesis estadística.

Pensemos en las consecuencias de nuestra elección antes de decidirnos. Si tenemos un valor para $p$, podemos calcular la probabilidad del suceso cuya observación es el dato esencial. Quiero significar la probabilidad de encontrar exactamente $m$ artículos defectuosos en una muestra azarosa de $n$ artículos. Llamemos a esta probabilidad $P$. Luego,

$$P = \binom{n}{m} p^m (1-p)^{n-m}.$$

El valor de $P$ depende de $p$, varía con ésta, pudiendo ser mayor o menor. No obstante, si esta probabilidad $P$ de un suceso observado es muy pequeña, rechazaríamos la hipótesis estadística subyacente. Sería tonto elegir una hipótesis tan inverosímil que ha de ser rechazada inmediatamente. Por tanto, elijamos la hipótesis menos inverosímil, aquella para la que el peligro de rechazo es menor. Es decir, elijamos el valor de $p$ para el cual $P$ sea *tan grande como es posible.*

Ahora bien, si $P$ es un máximo, el log $P$ es también un máximo, y, por tanto,

$$\frac{d \log P}{dp} = \frac{m}{p} - \frac{n - m}{1 - p} = 0.$$

Esta ecuación da

$$p = \frac{m}{n}.$$

Y así, tras algunas consideraciones, hacemos la elección que estábamos tentados a hacer desde el principio: como una razonable aproximación a $p$, la probabilidad subyacente, elegimos $m/n$, la frecuencia relativa observada.

Pero nuestra consideración no es un mero rodeo. Podemos aprender mucho de ella.

Empecemos por examinar el papel de $P$. $P$ es la probabilidad de un cierto suceso $S$ observado ($m$ defectuosos en una muestra de tamaño $n$). Esta probabilidad es calculada sobre la base de la hipótesis estadística $H_p$ de que $100p$ es el porcentaje de los defectuosos en el lote. La probabilidad $P$ varía con la hipótesis $H_p$ (con el valor de $p$). La más pequeña $P$, la menos aceptable y verosímil, parece $H_p$. Así, nos vemos conducidos a considerar $P$ como indicativa de la verosimilitud de la hipótesis. Este término «verosimilitud» ha sido introducido antes (en la sección 7 (5)), en el mismo sentido, pero ahora podemos ver más claramente las razones para su introducción.

Hagamos hincapié en que hemos elegido entre las varias hipótesis estadísticas admisibles $H_p$ (con $0 \leqq p \leqq 1$), para la cual $P$, la verosimilitud de $H_p$, es tan grande como es posible. Tras esta elección hay un principio, llamado adecuadamente el *principio de máxima verosimilitud*, que guía también al estadístico en otros casos, menos obvios que el nuestro.

## 9. Juzgando conjeturas no estadísticas

Consideremos varios ejemplos con el objeto de ilustrar la misma situación fundamental desde distintos ángulos.

(1) El otro día conocí a un tal señor Morgenstern. Este nombre no es muy usual, pero no me resultaba desconocido. Hubo un autor alemán Morgenstern por cuya música poética yo siento una gran inclinación. Y, ¡ah, sí!, mi primo, que vive en Atlanta, Georgia, recientemente

empezó a trabajar en las oficinas de Mark Morgenstern & Co., para la consulta de ingenieros.

Al principio no pensé en el señor Morgenstern. No obstante, al cabo de un rato oí que estaba en una oficina de ingeniería. Más tarde llegaron hasta mí otros retazos de información. Tuve ocasión de oir que el primer nombre de mi nuevo conocido es Mark, y que su lugar de trabajo es Atlanta, Georgia. En este punto es muy difícil creer que este señor Morgenstern no es el patrón de mi primo. En efecto, le pregunté directamente al señor Morgenstern, y así era.

Esta trivial historia es bastante instructiva. (Está basada, por cierto, en la experiencial real, pero los nombres están cambiados, naturalmente, así como otras circunstancias irrelevantes.) Que dos personas tengan exactamente el mismo apellido no es improbable, sobre todo si el nombre es muy común, como Jones o Smith. Es más improbable que dos personas diferentes tengan el mismo nombre y apellidos, especialmente cuando es un nombre poco común, tal como Mark Morgenstern. Que dos personas tengan la misma profesión, o la misma gran ciudad como residencia, no es improbable. Pero es muy improbable que dos personas diferentes tomadas al azar tengan el mismo nombre poco usual, la misma ciudad como residencia y la misma ocupación. Una casual coincidencia era difícil de creer, y así, mi conjetura sobre mi reciente conocimiento del señor Morgenstern era bastante razonable. Es más, era correcta, aunque esto no suponga nada dados los méritos del caso. Mi conjetura era razonable, defendible, justificable sobre la base de la probabilidad considerada. Aun en el caso de que mi conjetura hubiese sido incorrecta no tendría motivo para avergonzarme de ella.

En este ejemplo no se ha dado ningún valor numérico para la probabilidad decisivamente conectada con el problema, pero una tosca estimación no se podría obtener sin bastante dificultad.

(2) Dos amigos que se reúnen inesperadamente deciden escribir una tarjeta postal a un tercer amigo. Pero no están bastante seguros de la dirección. Ambos recuerdan la ciudad (se trata de París) y la calle (el boulevard Raspail), pero los dos dudan el número. «Espera —dijo uno de los amigos—, pensemos en el número sin hablar entre nosotros, y cada uno apuntará el número cuando crea haberlo hallado.» Esta propuesta fue aceptada y resultó que ambos apuntaron el mismo número: 79 boulevard Raspail. Pusieron esta dirección en la tarjeta postal, que llegó así al tercer amigo. La dirección era correcta.

Sin embargo, ¿cuál fue la razón para adoptar el número 79? Puesto que no hablaron entre sí, los dos amigos hicieron funcionar sus memorias independientemente. Ellos sabían que el boulevard Raspail era bastante

largo para tener edificios numerados hasta el 100. Por tanto, parece razonable asumir que la probabilidad para una coincidencia casual no es superior a 1/100. Esta probabilidad es pequeña, y así, la hipótesis de una coincidencia casual aparece como inverosímil. De aquí la confianza en el número 79.

(3) Según el estado de cuenta del Banco, el balance de mi cuenta era de 331,49 dólares al fin del pasado mes. Calculo mi balance en la misma fecha por medio de mis notas y encuentro la misma cantida d Tras este acuerdo de dos cálculos me quedo satisfecho en la opinión de que ambos son correctos. ¿Es esto cierto? De ningún modo. Aunque ambos cálculos llegaron al mismo resultado, el resultado podría estar equivocado, y el acuerdo ser debido a la casualidad. ¿Es esto verosímil?

La cantidad, expresada en centavos, es un número con cinco dígitos. Si el último dígito fuese elegido al azar podría haber sido 0, o 1, o 2, ..., u 8 tanto como 9, y así, la probabilidad de que el último dígito sea 9 es sólo 1/10. Lo mismo es cierto para cada una de las otras cifras. De hecho, si todas las cifras fuesen elegidas al azar, el número podría ser cualquiera de los siguientes:

$$000{,}00, \quad 000{,}01, \quad 000{,}02, \ldots 999{,}99.$$

Aquí tenemos, obviamente, 100.000 números. Si el ensamblaje de las cinco cifras, 33149, fue engendrado de alguna manera puramente azarosa, todos los ensamblajes podrían surgir igualmente. Y, como hay 100.000 tales ensamblajes, la probabilidad de que cualquier número dado por adelantado se produzca es

$$\frac{1}{100.000} = \left(\frac{1}{10}\right)^5 = 10^{-5}.$$

Ahora bien, $10^{-5} = 0{,}00001$ es una probabilidad muy pequeña. Si, intentando producir un efecto con tan pequeña probabilidad, alguien se las arregla para acertar a la primera, el resultado nos parecerá milagroso. No obstante, por mi parte me siento inclinado a creer que no hay nada milagroso en mi modesta cuenta del Banco. Una coincidencia casual es difícil de creer, y así, me veo conducido a la conclusión de que el acuerdo entre los dos cálculos se debe a la corrección del resultado. La gente normal generalmente piensa así en circunstancias semejantes, y tras las anteriores consideraciones esta clase de creencia parece bastante razonable.

(4) ¿Con qué lengua resulta el inglés más íntimamente relacionado, a cuál es más próximo, al húngaro o al polaco? Se necesita muy poco

conocimiento lingüístico para contestar esta pregunta, pero es más divertido obtener la contestación por nuestros propios medios que aceptar la autoridad de algún libro. El sentido común tiene un acceso a la cuestión.

Tanto la forma como el significado de las palabras cambian en el curso de la historia. Podemos comprender los cambios de forma si nos damos cuenta de que el mismo idioma es pronunciado de modo diferente en diferentes regiones, y podemos comprender los cambios de significado si nos damos cuenta de que los significados de las palabras no son fijos ni rígidos, sino flexibles y cambian con el contexto. No obstante, en este segundo aspecto hay una notable excepción: el significado de los números 1, 2, 3, ..., no puede variar ciertamente en grados perceptibles. Esta es una buena razón para sospechar que los números no cambian de significado a lo largo de la historia lingüística. Por tanto, basemos una primera comparación de las lenguas en la cuestión de los números. La tabla III contiene los diez primeros números en inglés, polaco, húngaro, y otras siete lenguas europeas modernas. Sólo consideramos las lenguas que utilizan el alfabeto latino (esto explica la ausencia del ruso y del griego moderno). Se omiten ciertos signos (acentos, cedillas) en sueco, alemán, polaco y húngaro, por ser desconocidos en español.

Mirando la tabla III y observando cómo los mismos números son escritos en diferentes idiomas, percibimos varias semejanzas y coincidencias. Los primeros cinco lenguajes (inglés, sueco, danés, holandés y alemán) parecen ser bastante semejantes unos a otros, y los próximos tres (francés, español e italiano) parecen estar en más íntimo acuerdo; así, tenemos dos grupos, uno que consta de cinco lenguajes y otro de tres. Aun así, estos dos grupos parecen estar íntimamente relacionados; observemos la ortografía coincidente de 3 en sueco, danés e italiano, o de 6 en inglés y francés.

TABLA III

LOS NUMEROS EN DIEZ IDIOMAS

| Inglés | Sueco | Danés | Holandés | Alemán | Francés | Español | Italiano | Polaco | Húngaro |
|---|---|---|---|---|---|---|---|---|---|
| one | en | en | een | ein | un | uno | uno | jedem | egy |
| two | tva | to | twee | zwei | deux | dos | due | dwa | ketto |
| three | tre | tre | drie | drei | trois | tres | tre | trzy | harom |
| four | fyra | fire | vier | vier | quatre | cuatro | quattro | cztery | negy |
| five | fem | fem | vijf | funf | cinq | cinco | cinque | piec | ot |
| six | sex | seks | zes | sechs | six | seis | sei | szesc | hat |
| seven | sju | syv | zeven | sieben | sept | siete | sette | siedem | het |
| eight | atta | otte | acht | acht | huit | ocho | otto | osiem | nyolc |
| nine | nio | ni | negen | neun | neuf | nueve | nove | dziewiec | kilenc |
| ten | tio | ti | tien | zehn | dix | diez | dieci | dziesiec | tiz |

El polaco parece más próximo a un grupo en algunos aspectos y a otro en otros; comparemos el 2 en sueco y en polaco, el 7 en español y en polaco. Sin embargo, el húngaro no muestra tales coincidencias con ninguna de las otras nueve lenguas. Estas observaciones dan la impresión de que esta lengua tiene poca relación con las otras, las cuales se relacionan todas entre sí de alguna manera. Especialmente, y ésta es la respuesta a nuestra pregunta inicial, el inglés parece estar definitivamente más estrechamente relacionado con el polaco que con el húngaro.

No obstante, hay varias objeciones. La primera es que «semejanza» y «acuerdo» son palabras vagas; diremos más precisamente lo que queremos expresar. Esta objeción apunta en la dirección correcta. Siguiendo su sugerencia, sacrificaremos una parte de nuestra evidencia para hacer la parte restante más precisa. Consideremos sólo las *iniciales* de los números apuntados en la tabla III. Comparemos dos números que expresan el mismo número en dos lenguas diferentes; los llamaremos «concordantes» si tienen la misma inicial, y «discordantes» si las iniciales son diferentes. La tabla IV contiene el número de casos concordantes por cada par de lenguas. Por ejemplo, el número 7, en la misma fila que las letras «Sp» y en la misma columna que la letra «P» indica que el español y el polaco tienen exactamente siete números concordantes de los diez casos posibles. El lector comprobará esta y otras cuantas concordancias

TABLA IV

INICIALES CONCORDANTES DE LOS NUMEROS EN DIEZ IDIOMAS

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| In | 8 | 8 | 3 | 4 | 4 | 4 | 4 | 3 | 1 | 39 |
| | S | 9 | 5 | 6 | 4 | 4 | 4 | 3 | 2 | 45 |
| | | D | 4 | 5 | 4 | 5 | 5 | 4 | 2 | 46 |
| | | | Ho | 5 | 1 | 1 | 1 | 0 | 2 | 22 |
| | | | | A | 3 | 3 | 3 | 2 | 1 | 32 |
| | | | | | F | 8 | 9 | 5 | 0 | 38 |
| | | | | | | Ep | 9 | 7 | 0 | 41 |
| | | | | | | | It | 6 | 0 | 41 |
| | | | | | | | | P | 0 | 30 |
| | | | | | | | | | Hu | 8 |

de la tabla IV. La última columna de la tabla IV nos muestra cuántos casos concordantes tiene cada lengua con las otras nueve lenguas juntas. Nos muestra también con bastante claridad la posición aislada del húngaro: éste tiene sólo ocho casos concordantes en total, mientras el número de casos concordantes de las otras nueve lenguas varía entre veintidós y cuarenta y seis.

Sin embargo, quizá, cualquier conclusión definitiva de estos datos

es precipitada: las citadas coincidencias de iniciales pueden ser debidas al azar. Es fácil proponer esta objeción, pero no tanto contestarla. El azar puede penetrar en el cuadro por varias vías. Puede haber un elemento de azar debido al hecho de que la correspondencia entre letras y pronunciación no es, de ninguna manera, rígida. Esto es cierto, incluso, en un solo lenguaje, especialmente en el inglés. *A fortiori*, la misma letra se pronuncia a menudo de modo diferente en diferentes idiomas y, por otro lado, se pronuncian de modo semejante diferentes letras. Tenemos que admitir que las coincidencias observadas no están libres de algún elemento azaroso. Pero la cuestión es la siguiente: ¿Es *probable* que tales coincidencias observadas sean debidas meramente al azar?

Si deseamos contestar esta pregunta con precisión, numéricamente, tenemos que aceptar alguna hipótesis estadística precisa y numéricamente definida y después sacar consecuencias de ella, que puedan ser confrontadas con las observaciones. Pero la elección de una hipótesis conveniente no resulta demasiado obvia. Consideremos aquí dos hipótesis estadísticas diferentes.

I. Tenemos dos sacos. Cada saco contiene veintiséis bolas, cada bola está señalada con una letra del alfabeto y bolas diferentes del mismo saco están marcadas de distinta manera. Con ambas manos sacamos simultáneamente de los dos sacos una bola de cada uno de ellos. Las dos letras así sacadas pueden o no coincidir; su coincidencia se asemeja a la coincidencia de las iniciales del mismo número escritas en dos lenguas diferentes (y la no coincidencia se asemeja a una no coincidencia entre tales iniciales). La probabilidad de una coincidencia es 1/26.

II. La coincidencia de las iniciales del mismo número escrito en dos idiomas distintos se asemeja de nuevo a la coincidencia de dos letras sacadas simultáneamente de dos sacos diferentes y, de nuevo, ambos sacos se llenan del mismo modo de bolas marcadas con letras. Pero ahora cada uno de los sacos contiene cien bolas y cada letra del alfabeto se usa para marcar tantas bolas diferentes del saco como números hay en la tabla III que tengan esa letra como inicial. La probabilidad de una coincidencia resulta ser 0,0948.

En ambas hipótesis la comparación de los diez primeros números se asemeja a diez extracciones independientes de la misma naturaleza.

Podemos comparar ambas hipótesis con las observaciones si calculamos probabilidades adecuadas. Las tablas V y VI contienen el material relevante.

La tabla V compara las frecuencias relativas realmente encontradas con las probabilidades calculadas. Las columnas (2) y (3) de la tabla V se refieren a los cuarenta y cinco pares de idiomas considerados en la

TABLA V

FRECUENCIAS ABSOLUTAS Y RELATIVAS, Y PROBABILIDADES, PARA $n$ O MAS COINCIDENCIAS DE INICIALES

| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|
| | Frecuencias | | | | Probabilidades | |
| | Diez idiomas | | Nueve idiomas versus húngaro | | Hipótesis II | Hipótesis I |
| $n$ | | | | | | |
| 0 | 45 | 1,000 | 9 | 1,000 | 1'000000 | 1,000000 |
| 1 | 40 | 0,889 | 5 | 0,556 | 0,630644 | 0,324436 |
| 2 | 35 | 0,778 | 3 | 0,333 | 0,243824 | 0,054210 |
| 3 | 31 | 0,689 | 0 | 0,000 | 0,061524 | 0,005569 |
| 4 | 25 | 0,556 | 0 | 0,000 | 0,010612 | 0,000381 |
| 5 | 15 | 0,333 | 0 | 0,000 | 0,001281 | 0,000018 |
| 6 | 9 | 0,200 | 0 | 0,000 | 0,000108 | 0,000001 |
| 7 | 7 | 0,156 | 0 | 0,000 | 0,000006 | 0,000000 |

tabla IV. Las columnas (4) y (5) de la tabla V se refieren sólo a nueve pares, formados por el húngaro emparejado con las nueve lenguas restantes. En gracia a la concreción fijemos nuestra atención en la línea que se ocupa de seis o más coincidencias ($n = 6$). Tales coincidencias se presentan en nueve casos de cuarenta y cinco, como muestra la columna (2). Por tanto, la frecuencia relativa observada de seis o más coincidencias es $9/45 = 0,2$, mientras todas estas coincidencias tienen sólo un poco más de una oportunidad entre diez mil para ocurrir según la hipótesis II, y sólo una oportunidad entre un millón según la hipótesis I; ver las columnas (6) y (7), respectivamente. Observaciones semejantes aplicables a otras líneas de la tabla V: lo que ha sido observado en realidad parece extremadamente improbable en ambas hipótesis, de modo que tenemos fuertes motivos para rechazarlas. Pero las columnas (4) y (5) presentan un aspecto diferente: las coincidencias observadas son improbables según la hipótesis I, pero parecen bastante

TABLA VI

NUMERO TOTAL DE COINCIDENCIAS, OBSERVADAS Y TEORICAS

(Hipótesis II)

| | COINCIDENCIAS | | DESVIACIONES | |
|---|---|---|---|---|
| | Observadas | Esperadas | Real | Modelo |
| Diez idiomas..... | 171 | 46,66 | 128,34 | 7,60 |
| Nueve idiomas V. húngaro. | 8 | 8,53 | —0,53 | 2,78 |

normales y usuales desde el punto de vista de la hipótesis II. La impresión alcanzada por la tabla V es corroborada por la tabla VI: si consideramos los cuarenta y cinco pares de idiomas, el número total de coincidencias observado realmente excede muchísimo lo que habíamos esperado sobre la base de la hipótesis II, pero los números esperados y observados están muy de acuerdo si consideramos sólo los nueve pares en que el húngaro es emparejado con las otras nueve lenguas. (Con respecto a la hipótesis I, hay un desacuerdo mucho mayor en ambos casos.)

En resumen, no hay interpretación obvia del «azar» que nos permita dar cuenta de todas las coincidencias observables en la tabla III; hay demasiadas. Pero podemos, de modo bastante razonable, dar cuenta de las coincidencias entre el húngaro y las otras lenguas. La explicación de que el húngaro está separado de las otras lenguas que se relacionan entre sí resulta justificada.

El nudo está en que esta explicación ha sido justificada, gracias a la consideración de probabilidades, por unas *pocas observaciones*. La misma explicación es apoyada por un abrumador aparato de evidencia filológica.

(5) A partir de observaciones apropiadas (con telescopio y espectroscopio) podemos concluir que ciertos elementos encontrados en la corteza terrestre se hallan también en el Sol y en ciertos astros. Esta conclusión está basada en una ley física descubierta por G. Kirchhoff hace casi un siglo (que dice, *a grosso modo*, que un vapor luminoso absorbe exactamente la misma clase de luz que emite). No obstante, la conclusión apela también a las probabilidades, y éste es el punto que aquí nos interesa; reduciremos la parte física del argumento a un bosquejo esquemático.

Usando un aparato adecuado (un prisma o una difracción reticulada) podemos detectar una secuencia de líneas en la luz del Sol (en el espectro solar). Podemos detectar también una secuencia de líneas en la luz emitida por ciertas sustancias, tales como el hierro evaporado a alta temperatura en el laboratorio. (En efecto, las líneas en el espectro del Sol, líneas Fraunhofer, son opacas y las líneas del espectro del hierro son brillantes). Kirchhoff examinó sesenta líneas de hierro y encontró que cada una de estas líneas coincide con alguna línea solar. (Ver el tosco bosquejo de la figura 14.5 o la decimocuarta edición de la *Encyclopaedia Britannica*, volumen 21, figura 3, grabado I, vuelta de la página 560.) Estas coincidencias son totalmente comprensibles si damos por supuesto que hay hierro en el Sol. (Más exactamente, estas consecuencias se siguen de la ley de Kirchhoff sobre la emisión y absorción si suponemos que en la atmósfera del Sol hay vapor de hierro que absorbe

algo de la luz emitida por la parte central del Sol incandescente a temperaturas aún más altas.) No obstante, quizá, y aquí está de nuevo la objeción siempre presente, estas coincidencias se deben al azar.

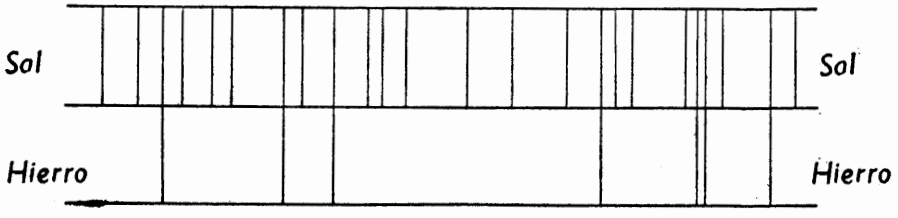



Fig. 14.5. Coincidencias.

La objeción merece seria consideración. En efecto, ninguna observación física es absolutamente precisa. Dos líneas que se nos muestran coincidentes podrían ser diferentes en realidad y, sólo por azar, estar tan unidas la una a la otra que, con la limitada precisión de nuestras observaciones, no pudiéramos reconocer su diferencia. Tenemos que conceder que cualquier coincidencia observada puede ser sólo una coincidencia aparente, y que puede haber, en efecto, una pequeña diferencia. No obstante, hagámonos una pregunta: ¿Es *probable* que cada una de las sesenta coincidencias observadas se produzcan por diferencias azarosas tan pequeñas que no logremos detectarlas por los medios de observación empleados?

Kirchhoff, que registró las líneas observadas en una escala (arbitraria) de centímetros, estimó que no podía haber fracasado en reconocer una diferencia que excediese medio milímetro de su escala. En esta escala la distancia promedio entre dos líneas adyacentes del espectro solar era de unos dos milímetros. Si las sesenta líneas de hierro fuesen arrojadas al azar en este cuadro, independientemente unas de otras, ¿cuál sería la probabilidad de que cada una cayese a menos de medio milímetro de distancia de alguna línea solar?

Encaminaremos esta pregunta a su solución formulando otra equivalente en un dominio más familiar. Algunas líneas paralelas están pintadas en el suelo; la distancia media entre dos líneas adyacentes es de dos pulgadas. Arrojamos una moneda al suelo sesenta veces. Si el diámetro de la moneda es de una pulgada ¿cuál será la probabilidad de que la moneda cubra una línea cada vez?

En esta última formulación la pregunta es fácil de responder. Supongamos primero que las líneas son equidistantes (como en la figura 14.6) de manera que la distancia de cada línea a la siguiente es de dos pulgadas. Si la moneda cubre una línea, su centro (el de la moneda)

está a lo sumo a media pulgada de distancia de la línea y, por tanto, este centro se encuentra en una faja de media pulgada de ancha dividida en dos por la línea (sombreada en la figura 14.6). Como es obvio, la probabilidad de que la moneda arrojada al suelo cubra una línea es de 1/2. La probabilidad de que la moneda, arrojada al suelo sesenta veces, cubriese alguna línea cada vez es de $(1/2)^{60}$.

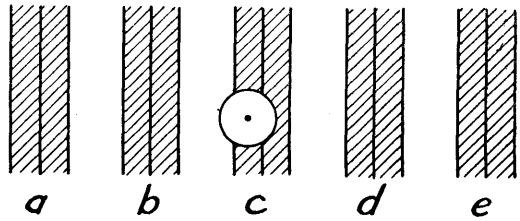



Fig. 14.6. Líneas equidistantes.

Supongamos ahora que las líneas del suelo no son equidistantes; la distancia media entre dos líneas adyacentes se supone que sea aún de dos pulgadas. Imaginemos que las líneas que fueron originalmente equidistantes ocupan su posición presente por haber sido movidas sucesivamente. Si una línea (como la línea *b* de la figura 14.7) se mueve de modo que su distancia a la más próxima resulta ser más de una pulgada, las posibilidades que tiene la moneda de cubrir alguna línea no han cambiado. Si, no obstante, la línea se ha movido (como la línea *d* de la figura 14.7) tanto que su distancia a la próxima es menor de una pulgada, las dos fajas unidas (sombreadas) quedan sobrepuestas y las posibilidades de la moneda de cubrir una línea disminuyen. Por tanto, la probabilidad requerida es menor de $(1/2)^{60}$.

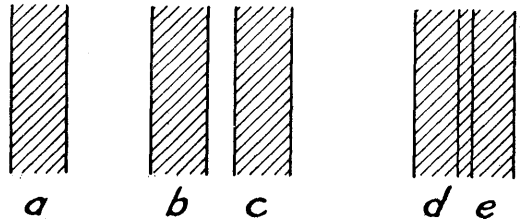



Fig. 14.7. Líneas a distancias irregulares.

Para resumir: si las líneas de hierro fueran arrojadas por un ciego, al azar, en el espectro solar la probabilidad de las sesenta coincidencias observadas por Kirchhoff sería menor de $2^{-60}$ y, así, menor de $10^{-18}$, o

$$\frac{1}{1.000.000.000.000.000.000}.$$

«Esta probabilidad —dice Kirchhoff, a quien citamos ya en el lema con que iniciamos el capítulo— se hace aún más pequeña por el hecho de que cuanto más brillante vemos que es una línea de hierro determinada más opaca aparece la correspondiente línea solar. De aquí que esta coincidencia debe ser producida por alguna causa, y se puede asignar una causa que proporcione una explicación perfecta de los hechos observados.»

(6) El siguiente ejemplo no se basa en la observación real, pero ilustra una importante situación que surge con frecuencia.

Una enfermedad extremadamente peligrosa ha sido tratada en la misma localidad por dos métodos diferentes, a los que llamaremos «viejo tratamiento» y «nuevo tratamiento». De los nueve pacientes a quienes se dio el «viejo tratamiento» seis murieron y sólo tres sobrevivieron, mientras que de once pacientes que recibieron el «nuevo tratamiento» sólo dos murieron y nueve sobrevivieron. Los veinte casos están claramente expuestos en la tabla VII.

TABLA VII

TABLA DE CUATRO CORRELACIONES

| PACIENTES | Muertos | Sobre-vivientes | TOTAL |
|---|---|---|---|
| «Viejo tratamiento»..... | 6 | 3 | 9 |
| «Nuevo tratamiento»..... | 2 | 9 | 11 |
| TOTAL..... | 8 | 12 | 20 |

Una primera ojeada a esta tabla nos da la impresión de que las observaciones apuntadas hablan muy fuertemente a favor del «nuevo tratamiento». La frecuencia relativa de casos fatales es

6/9, o 67 por 100, con el «viejo tratamiento»,

2/11, o 18 por 100, con el «nuevo tratamiento».

No obstante, pensándolo mejor, podemos preguntarnos si los números observados son bastante grandes para darnos un grado de confianza razonable en los porcentajes calculados, 67 y 18 por 100. Aun así, queda el hecho de que el número de casos fatales fue mucho más bajo con el «nuevo tratamiento». Sin embargo, tan baja mortalidad puede ser debida al azar. *¿Hasta qué punto puede el azar producir tal resultado?*

Esta última pregunta parece ser la correcta. De todas formas, la pregunta debe ser expuesta con más precisión antes de contestarla. Tenemos que explicar el significado preciso que damos a las palabras «azar» y «tal». La primera de ellas quedará explicada si asimilamos el caso presente a algún juego de azar conveniente. Una buena interpretación de las palabras «tal resultado» puede ser la siguiente: Consideremos todos los resultados en que el número de desgracias con el segundo tratamiento *no es más alto que el realmente observado.* Así, podemos llegar, al fin, a la formulación siguiente:

*Hay dos jugadores, el señor Oldman y el señor Newman, y veinte cartas, de las cuales ocho son negras y doce rojas. Las cartas se reparten de modo que el señor Oldman recibe nueve y el señor Newnam once. ¿Cuál es la probabilidad de que el señor Newman reciba dos o menos cartas negras?*

Esta formulación expresa del modo más sencillo y agudo posible el contenido que hemos de examinar: la diferencia entre el «viejo» y el «nuevo tratamiento» no influye realmente en la mortandad, y el resultado observado se debe meramente al azar.

La probabilidad requerida resulta ser

$$\frac{335}{8398} = 0,0399 \sim \frac{1}{25}.$$

Es decir, un resultado que parece tan favorable al «nuevo tratamiento» como el resultado observado, o aún más, se producirá por casualidad una vez, aproximadamente, en veinticinco ensayos. Y así, la evidencia numérica de la superioridad del «nuevo tratamiento» no puede ser sencillamente abandonada, pero, ciertamente, no es muy grande.

Para juzgar con claridad en estas cuestiones vamos a considerar por un momento una situación en que los datos numéricos nos lleven a una probabilidad de 1/10.000 en vez de 1/25. Tales datos harían muy difícil creer que la diferencia observada en la mortalidad se debe meramente al azar, pero, naturalmente, no prueban inmediatamente la superioridad del «nuevo tratamiento». Los datos proporcionarían un argumento bastante fuerte en favor de la existencia de alguna *diferencia no fortuita* entre las dos clases de casos. Los números no pueden decir cuál es la naturaleza de esta diferencia realmente; si sólo recibieran el «nuevo tratamiento» personas jóvenes o vigorosas, y sólo recibiera el «viejo» gente anciana o débil, el argumento en favor de la superioridad médica de uno de los tratamientos sería extremadamente flojo.

(7) Creo que el lector habrá observado cierto paralelismo entre los seis ejemplos precedentes de esta sección. Ahora este paralelismo puede

estar dispuesto para ser formulado en términos generales. Sin embargo, sigamos cuanto podamos el ejemplo del naturalista que compara cuidadosamente los detalles importantes más que el de aquellos filósofos que confían principalmente en las palabras. Al examinar nuestros ejemplos observamos bastantes detalles; si no tenemos en cuenta los pormenores relevantes se habrá perdido nuestra labor.

En cada ejemplo hay una *coincidencia* y una *explicación.* (El nombre, apellido, ocupación y ciudad natal de una persona que conozco coinciden con los de una persona de quien he oído hablar. Explicación: las dos personas son la misma.—Dos números recordados o calculados por dos diferentes personas coinciden. Explicación: el número al que han llegado dos personas, trabajando independientemente, es correcto.— Las iniciales de varios pares de números, que designan la misma cifra en dos lenguajes diferentes, coinciden. Explicación: los lenguajes están relacionados.— Las líneas brillantes del espectro del hierro, observadas en los experimentos del laboratorio, coinciden con ciertas líneas opacas del espectro del Sol. Explicación: hay vapor de hierro en la atmósfera del Sol.— Un nuevo tratamiento de una enfermedad coincide con una más baja mortandad. Explicación: el nuevo tratamiento es más efectivo.)

En contraste con estas explicaciones específicas, cuya naturaleza varía con la naturaleza del ejemplo, hay otra explicación que puede ser establecida en los mismos términos en todos los ejemplos: las coincidencias observadas son debidas al azar.

Las explicaciones específicas no carecen de fundamento, algunas de ellas son convincentes, pero ninguna de ellas es necesaria lógicamente o rigurosamente comprobada. Por tanto, la situación es fundamentalmente la misma en cada ejemplo: hay dos conjeturas rivales, una conjetura específica y la «universalmente aplicable» hipótesis de lo «azaroso» que atribuye las coincidencias al azar.

Sin embargo, si observamos más atentamente, notaremos que la «hipótesis de lo azaroso» es vaga. El enunciado «ese efecto se debe al azar» es ambiguo, puesto que el azar puede actuar según esquemas diferentes. Si deseamos obtener alguna indicación más definida de ello tenemos que hacer la hipótesis de lo azaroso más precisa, más específica y expresarla en términos de probabilidad; en resumen, tenemos que elevarla a la categoría de una *hipótesis estadística.*

En los asuntos cotidianos no nos tomamos la molestia de establecer una hipótesis estadística con precisión o de calcular numéricamente su verosimilitud. No obstante, podemos dar un primer paso en esta dirección (como en el ejemplo (1)) o incluso ir un poco más lejos (como

en los ejemplos (2) o (3)). En las cuestiones científicas formularíamos claramente la hipótesis estadística implicada y continuaríamos con un estimado numérico de su verosimilitud, como en los ejemplos (5) y(6).

En la transición de la idea general y, por tanto, algo difusa del azar, a una hipótesis estadística específica tenemos que hacer una elección. Hay casos en que apenas nos damos cuenta de esta elección, puesto que sólo percibimos que es bastante sencilla y se adecua razonablemente bien al caso; en un caso semejante la hipótesis elegida parece «natural» (como en los ejemplos (3), (5) y (6)). En otros casos la elección es bastante perceptible; no vemos inmediatamente una hipótesis estadística que sería bastante sencilla y cumpliría el caso «realísticamente»; así, después de más o menos dudas, elegimos (como en el ejemplo (4)).

Al final tenemos dos conjeturas rivales una frente a la otra: una conjetura no estadística, digamos «física», que es la conjetura $F$, y una hipótesis estadística $E$. Ahora bien, se ha observado cierto acontecimiento $A$. Este acontecimiento $A$ se relaciona con ambas conjeturas $F$ y $E$, y está tan relacionado con ella que podría influir en nuestra elección entre las dos conjeturas rivales $F$ y $E$, si sucediera. Si la conjetura física $F$ es verdadera, $A$ parece fácilmente explicable y su acaecer es muy comprensible. En los casos más claros (como en el ejemplo (5)) $A$ está implicada en $F$, es una consecuencia suya. Por otro lado, desde el punto de vista de la hipótesis estadística $E$, el acontecimiento $A$ aparece como una «coincidencia», cuya probabilidad $p$ puede ser calculada sobre la base de la hipótesis $E$. Si la probabilidad $p$ de $A$ resulta baja, el acaecer del suceso $A$ no se explica fácilmente por «azar», esto es, por la hipótesis estadística $E$; esto debilita nuestra confianza en $E$ y refuerza nuestra confianza en $F$. Por otra parte, si la probabilidad $p$ del acontecimiento observado $A$ es alta, $A$ puede parecer explicable por azar, esto es, por la hipótesis estadística $E$; nuestra confianza en $E$ resulta reforzada y debilitada nuestra confianza en $F$.

Observaremos que lo anterior está de acuerdo con lo que decimos sobre conjeturas rivales en la sección 13.12 y añade alguna precisión al patrón de razonamiento plausible examinado en la sección 12.3.

La omnipotente hipótesis de lo azaroso es una alternativa a cualquier otra clase de explicaciones; parece estar profundamente enraizada en la naturaleza humana. «¿Ocurrió intencionadamente o fue un accidente?» «¿Existe una causa determinable o es una coincidencia meramente casual?» En casi cada debate o deliberación, en el trivial chismorreo y en los tribunales, en los asuntos cotidianos y en la ciencia, se dan preguntas de esta clase.

## 10. Juzgando conjeturas matemáticas

Comparemos algunos ejemplos, tratados en los anteriores capítulos, entre sí y con los tratados en la sección anterior.

(1) Recordemos la historia de un notable descubrimiento referido en la sección 2.6. Euler investigó la serie infinita

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \ldots + \frac{1}{n^2} + \ldots$$

Primero encontró varias transformaciones de esta serie. Luego, utilizando una de estas transformaciones, obtuvo un valor numérico aproximado para la suma de la serie, el valor 1.644934. Finalmente, por un nuevo y osado procedimiento, intuyó que la suma de la serie es $\pi^2/6$. El mismo Euler se dio cuenta de que este procedimiento era atrevido e incluso objetable, pero tenía una buena razón para confiar en su descubrimiento: el valor encontrado por cálculo numérico, 1.644934, coincidía, hasta donde él llegó, con el valor intuido

$$\frac{\pi^2}{6} = 1.64493406 \ldots$$

Y así, Euler confió en su intuición. Sin embargo, ¿era esta confianza razonable? Tal coincidencia podía deberse al azar.

En efecto, no es completamente imposible que tal coincidencia sea debida al azar, aunque hay sólo una oportunidad entre diez millones de que tal coincidencia ocurra: la probabilidad de que el azar, interpretado de un modo sencillo, produzca tal coincidencia de siete decimales es $10^{-7}$; cf. sección 9 (3) y ejemplo 11. Así, pues, no podemos quejarnos de que Euler rechazase la explicación por coincidencia casual y se adhiriese a su intuición $\pi^2/6$. Al final demostró su intuición. Sin embargo, no necesitamos insistir aquí en el hecho de que la probase: con o sin confirmación, la intuición de Euler era, en sí misma, no sólo brillante, sino también razonable.

(2) Miremos la sección 3.1 de nuevo, especialmente en la figura 3.1, que presenta nueve poliedros. Por cada poliedro determinamos $C$, $V$ y $A$, es decir, el número de caras, vértices y aristas, respectivamente, y enumeramos los números encontrados en una tabla (pág. 66). Luego observamos una regularidad: a lo largo de toda la tabla.

$$C + V = A + 2.$$

Nos pareció improbable que una regularidad tan persistente fuese mera coincidencia, y así fuimos llevados a la conjetura de que la relación observada en nueve casos era cierta, en general.

Hay un punto en este razonamiento que puede hacerse más preciso: ¿Cuál es la probabilidad de tal coincidencia? Para contestar a esta pregunta tenemos que proponer una hipótesis estadística definida. Por mi parte no he sido capaz de pensar una que se ajuste al caso perfectamente, pero hay una que tiene algún sentido en esta situación. Permítaseme enunciarla así: $C - 1 = X$, $V - 1 = Y$, $A = Z$. Con este cambio de notación la relación conjetural obtiene la forma $X + Y = Z$.

Tenemos tres sacos, cada uno de las cuales contiene $n$ bolas, numeradas 1, 2, 3, ..., $n$. Saquemos una bola de cada saco y señalemos con $X$, $Y$ y $Z$ el número del primero, del segundo y del tercer saco, respectivamente. ¿Cuál es la probabilidad de que encontremos la relación

$$X + Y = Z$$

entre los tres números $X$, $Y$ y $Z$, producida por azar?

Se entiende que las tres extracciones son mutuamente independientes. Con esta condición es determinada la probabilidad requerida y nos resulta fácil encontrar que es igual a

$$\frac{n - 1}{2n^2}.$$

Apliquemos esto a nuestro ejemplo. Fijémonos en el momento en que logramos fijar la relación hipotética de un nuevo poliedro. Por ejemplo, después de los nueve poliedros que examinamos inicialmente (en la sección 3.1) tomamos el caso del icosaedro (en la sección 3.2). Para el icosaedro, como vimos, $C = 20$, $V = 12$, $A = 30$, etc.; en efecto,

$$(C - 1) + (V - 1) = 19 + 11 = 30 = A.$$

¿Es esto una coincidencia meramente azarosa? Apliquemos nuestra fórmula, tomando $n = 30$ (ciertamente, no podríamos hacer $n$ menor de 30), y encontremos que tal suceso tiene la probabilidad

$$\frac{29}{2 \times 30^2} = \frac{29}{1800} = 0{,}016111;$$

es decir, tiene algo menos de una oportunidad entre 60. Podemos dudar si adscribimos, o no, la verificación de la relación conjeturada a mero azar. Pero si logramos verificarla para otro poliedro con $C$, $V$ y $A$,

aproximadamento, tan grandes como para el icosaedro, y nos sentimos inclinados a mirar las dos verificaciones como mutuamente independientes, enfrentamos un suceso (la verificación conjunta de ambos casos) con una probabilidad menor de $(1/60)^2$; este suceso tiene menos oportunidad de ocurrir que una entre 3.600, y es, por tanto, más difícil de explicar por azar. Si las verificaciones continúan sin interrupción, llega un momento, más pronto o más tarde, que nos sentimos inclinados a rechazar la explicación por azar.

(3) En el anterior ejemplo no hemos puesto demasiada atención sobre los valores numéricos de las probabilidades calculadas. Para darnos cuenta de que la probabilidad disminuye, a medida que una verificación sigue a otra, puede ser más adecuada la orientación de nuestro juicio que los valores numéricos calculados. De cualquier modo, hay casos en que sería difícil ofrecer una hipótesis estadística adecuada y, así, resulta imposible calcular las probabilidades envueltas numéricamente, aunque el cálculo de probabilidad aún nos dé sugestiones valiosas.

En la sección 4.8 hemos comparado dos conjeturas referentes a la suma de dos cuadrados. Llamémolas conjeturas $A$ y $B$, respectivamente. La conjetura $A$ (que hemos descubierto al final de la sección 4.6) propone una notable regla que determina precisamente de cuántas maneras un entero de cierta forma puede ser representado como una suma de cuatro cuadrados impares. La conjetura $B$ (conjetura de Bachet) afirma que cualquier entero puede ser representado como la suma de cuatro cuadrados de una o más maneras. Las dos conjeturas ofrecen una predicción sobre la suma de cuadro cuadrados, pero la predicción presentada por $A$ es más precisa que la presentada por $B$. Sólo para acentuar este punto consideremos un momento una suposición bastante increíble. Asumamos que (de fuente misteriosa) sabemos que, en un caso determinado, el número de representaciones tiene una oportunidad igual para cualquiera de los valores $r + 1$, esto es, 0, 1, 2, ... $r$, y no puede tener un valor que exceda de $r$, que es un número bastante grande (y esto mantiene ambas bajo las circunstancias especificadas en $A$ y bajo las especificadas en $B$ —una asunción bastante absurda—). Ahora bien, $A$ predice que el número de representaciones tiene un valor definido; $B$ predice que es más grande que 0. Por tanto, la probabilidad de que $A$ sea cierta en el caso asumido es $1/(r + 1)$, mientras la probabilidad de que $B$ sea cierta es $r/(r + 1)$. De hecho, tanto $A$ como $B$ son ciertas en ese caso, las dos conjeturas se verifican y la pregunta es entonces qué verificación da la evidencia más fuerte. A la vista de lo que hemos examinado es mucho más difícil atribuir la verificación de $A$ al azar que la verificación de $B$. En virtud de esta circuns-

tancia (de acuerdo con los casos semejantes examinados en este capítulo), la verificación de la predicción más precisa $A$ tendrá más peso que la verificación de la menos precisa $B$. En la sección 4.8 llegamos al mismo punto de vista sin ninguna consideración explícita de probabilidades.

## EJEMPLOS Y COMENTARIOS

### Primera parte

Cada ejemplo de esta primera parte empieza con una referencia a alguna sección o subsección de este capítulo y da fórmulas o derivaciones omitidas en el texto. Las soluciones requieren algún conocimiento del cálculo de probabilidades.

**1.** [Sección 3 (3).] Aceptar el esquema de la sección 3 (3) para representar la sucesión de los días «lluviosos» y «sin lluvia». Decir «soleado» en vez de «sin lluvia» en gracia a la conveniencia, y denotemos con $r_r$, $s_r$, $r_s$ y $s_s$ varias probabilidades,

$r_r$ para un día lluvioso después de un día lluvioso,
$s_r$ para un día soleado después de un día lluvioso,
$r_s$ para un día lluvioso después de un día soleado, y
$s_s$ para un día soleado después de un día soleado.

(*a*) Muéstrese que $r_r - r_s = s_s - s_r$.

(*b*) Se dice que «a un día lluvioso sigue un día lluvioso con más facilidad que uno sin lluvia». ¿Qué significa esto exactamente?

**2.** [Sección 3 (4).] Se dice que cada carta tiende a ser diferente de la anterior. ¿Qué significa esto exactamente?

**3.** [Sección 5 (1).] Encontrar la expresión general para los números de la columna (3) de la tabla I.

**4.** [Sección 5 (2).] Encontrar la expresión general para los números de la columna (5) de la tabla I.

**5.** [Sección 5 (3).] (*a*) Encontrar la expresión general para los números de la columna (3) de la tabla II.

(*b*) Al objeto de detectar una desviación sistemática, si hay alguna, examinar las diferencias de los datos correspondientes (en la misma fila) de las columnas (4) y (5); enumerar los signos.

**6.** [Sección 7 (1).] Si un ensayo consiste en arrojar tres dados «justos» y un éxito consiste en sacar seis puntos con cada dado, ¿cuál es la probabilidad de $n$ éxitos en $n$ ensayos?

**7.** [Sección 7 (2).] Entre los varios sucesos de la historia del reverendo Galiani contada en la sección 7 (1), ¿cuál constituye el argumento más fuerte contra la hipótesis del dado «justo»?

**8.** [Sección 7 (3).] (*a*) Escribir la fórmula que conduce al valor numérico $1{,}983 \cdot 10^{-7}$.

(*b*) La probabilidad de un éxito es 1.3. Encontrar la probabilidad de que 315672 ensayos den exactamente 315672/3 éxitos.

**9.** [Sección 8 (1).] La expresión dada por *a* es una suma. Cada término de esta suma es una probabilidad: ¿Para qué?

**10.** [Sección 8 (1).] Encontrar la abscisa del punto de inflexión de la curva representada por la figura 14.4.

**11.** [Sección 9 (3).] Dado un número de *n* cifras. Se produce al azar una secuencia de *n* cifras, quizá por un mono jugando con las teclas de una máquina de sumar. ¿Cuál es la probabilidad de que la secuencia así producida coincida con el número dado? ¿Es matemáticamente determinada la contestación?

**12.** [Sección 9 (4).] Explicar el cálculo de la probabilidad 0,0948.

**13.** [Sección 9 (4).] Encontrar la expresión general para los números: (*a*) de la columna (6), (*b*) de la columna (7), de la tabla V.

**14.** [Sección 9 (4).] Explicar el cálculo de los números de coincidencias esperados en la tabla VI: (*a*) 42,66, (*b*) 8,53.

**15.** [Sección 9 (4).] Explicar el cálculo de la desviación modelo 2,78 en la última fila y la última columna de la tabla VI.

**16.** [Sección 9 (5).] ¿Por qué $(1/2)^{60}$?

**17.** [Sección 9 (6).] Explicar el cálculo de la probabilidad 0,0399. Generalizar.

**18.** [Sección 10 (2).] Derivar la expresión $(n-1)/2n^2$ para la probabilidad requerida.

## Segunda parte

**19.** *Sobre el concepto de probabilidad.* La sección 2 no define lo que «es» la probabilidad; simplemente intenta explicar lo que apunta a describir: la frecuencia relativa «a largo alcance», la frecuencia relativa de «final fijo» o la frecuencia relativa en una «muy larga» serie de observaciones. Cuán larga sea la serie no queda establecido. Esto es una omisión.

Pero tales omisiones son frecuentes en ciencia. Tomemos la parte más vieja de la física, la mecánica, y veamos la definición de velocidad rectilínea de movimiento no uniforme: la velocidad es el espacio descrito por el punto que se mueve en un cierto intervalo de tiempo, dividido por la longitud de ese intervalo, siempre que el intervalo sea «muy corto». Cuán corto se supone que sea ese intervalo tampoco está establecido.

En la práctica nosotros tomamos el intervalo de tiempo tan corto como nos permitan nuestros medios de observación, o en el caso de la serie estadística observada tan larga como nos permitan los mismos. Teóricamente podemos pasar al límite. Los físicos, al definir la velocidad, suponen que el intervalo de tiempo tiende a cero. R. von Mises supone, al definir la probabilidad, que la serie estadística tiende a infinito.

**20.** *Cómo no debe ser interpretado el concepto de frecuencia de la probabilidad.* El Dr. Tel. movió su cabeza cuando terminó de reconocer al paciente. (Dr. Tel. significa doctor en telepatía; aunque contando con la denonada oposición de la profesión médica, la práctica de la telepatía ha sido reconocida en los cincuenta y tres Estados de la Unión.) «Usted tiene una enfermedad muy seria —dijo el Dr. Tel—. De diez personas que cogen esta enfermedad sólo una sobrevive.» Cuando el paciente estaba muy asustado por la noticia, el Dr. Tel. continuó: «Pero usted tiene mucha suerte. Usted sobrevivirá, puesto que vino a mí. Yo ya he tenido nueve pacientes que han muerto de ella».

Quizá el Dr. Tel. lo creía así. Su abuelo era un marinero cuyo barco fue golpeado por un molusco en una salida al mar. El marinero puso su cabeza en el agujero hecho por el molusco en el casco del barco y así se sintió protegido «porque —razonó— es muy improbable que el mismo molusco golpee el mismo lugar dos veces».

**21.** Un oficial, encargado de supervisar una elección en cierta localidad, encontró treinta registros engañosos entre los treinta y ocho que examinó la primera mañana. Un periódico declaró que por lo menos el 99 por 100 de los registros de la localidad eran correctos y más allá de toda sospecha. ¿Cómo juzgar la afirmación del diario a la luz de la observación del oficial?

**22.** En el escaparate de un relojero hay cuatro «cucús», funcionando. Tres de los cuatro relojes están separados por menos de dos minutos: ¿Podemos confiar en la hora que tienen? Hay una conjetura natural: los relojes fueron puestos en hora al principio, aunque no muy precisamente (se trataba de «cucús») y uno de ellos no funciona. Si es así, podemos confiar en la hora de los tres restantes. Pero hay una conjetura

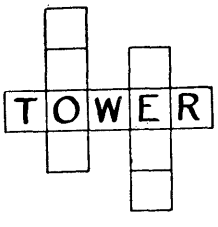
rival, naturalmente: los tres relojes van de acuerdo por casualidad. ¿Cuál es la probabilidad de que ocurra esto último?

**23.** Si $a$, $b$, $c$, $d$, $e$ y $f$ son enteros elegidos al azar, que no exceden en valor absoluto a un entero positivo dado $n$, ¿cuál es la probabilidad

$$ax + by = e$$
$$cx + dy = f$$

de que el sistema de dos ecuaciones con dos incógnitas tenga una sola solución?

**24.** *La probabilidad y la solución de problemas.* En un crucigrama una palabra desconocida de cinco letras se cruza con dos desconocidas palabras de cuatro letras cada una. Intuimos que la palabra desconocida de cinco letras es «TOWER»*, y así tendremos la situación indicada por el siguiente diagrama:




Al objeto de comprobar la intuición nos gustaría encontrar una o las otras cuatro palabras de cuatro letras que cruzan la conjetural TOWER. Una de las palabras cruzadas podría verificar la O, la otra la E. ¿Qué verificación daría más peso? Y ¿por qué?

**25.** *Regular o irregular.* Compárense las dos columnas de números:

| I | II |
|---|---|
| 1005 | 1004 |
| 1033 | 1038 |
| 1075 | 1072 |
| 1106 | 1106 |
| 1132 | 1139 |
| 1179 | 1173 |
| 1205 | 1206 |
| 1231 | 1239 |
| 1274 | 1271 |
| 1301 | 1303 |

* Por conveniencia hemos conservado la palabra inglesa «tower», aun a pesar de que la española equivalente («torre») tiene el mismo número de letras.—*N. del T.*

Una de estas dos columnas es «regular», la otra «irregular». La columna regular contiene diez mantisas sucesivas de una tabla de logaritmos vulgares de cuatro cifras. Los números de la columna irregular están de acuerdo con los números correspondientes de la columna regular en los tres primeros dígitos, pero el cuarto podría ser obra de un calculador poco confiable: éstos han sido elegidos «al azar». ¿Cuál es cuál? [Hallar un procedimiento ordenado para distinguir la regular de la irregular.]

**26.** *Las reglas fundamentales del cálculo de probabilidad.* Al calcular las probabilidades podemos visualizar el conjunto de casos posibles y ver intuitivamente que ninguno resulta privilegiado entre ellos, o podemos proceder según las reglas. Es importante para el principiante comprobar que puede llegar al mismo resultado por dos vías diferentes. Las reglas son particularmente importantes cuando miramos la teoría de la probabilidad como una teoría puramente matemática. Las reglas serán importantes en el próximo capítulo. Por estas razones, introduzcamos aquí las reglas fundamentales del cálculo de probabilidad usando el saco y las bolas*; cf. la sección 3.

El saco contiene $p$ bolas. Algunas de las bolas están marcadas con una $A$, otras con una $B$, algunas con ambas letras, y, por último, algunas no están marcadas de ningún modo. (Hay $p$ casos posibles y dos «propiedades» o «sucesos», $A$ y $B$.) Escribamos $A$ por ausencia de $\bar{A}$ o «no-A». (Tomamos —, el signo de la negación, pero lo colocamos arriba de la letra, no antes de ella.) Hay cuatro posibilidades, cuatro clases de bolas.

La bola tiene $A$, pero no tiene $B$. Denotamos esta categoría por $A\bar{B}$ y el número de bolas semejantes por $a$.

La bola tiene $B$, pero no tiene $A$. Denotamos esta categoría por $\bar{A}B$ y el número de bolas de esta clase por $b$.

La bola tiene $A$ y $B$. Denotamos esta categoría por $AB$ y el número de bolas tales (común a $A$ y $B$) por $c$.

La bola no tiene ni $A$ ni $B$. Denotamos esta categoría por $\bar{A}\bar{B}$ y el número de bolas (diferentes de la que tienen $A$ o $B$) por $d$.

Por tanto, obviamente,

$$a + b + c + d = p.$$

---

* Seguimos aquí a H. POINCARÉ, *Calcul des probabilités*, págs. 35-39.

$\Pr\{A\}$ es una abreviación para la probabilidad de $A$, y $\Pr\{B\}$ para la de $B$. Con esta notación tenemos

$$\Pr\{A\} = \frac{a + c}{p}, \qquad \Pr\{B\} = \frac{b + c}{p}.$$

$\Pr\{AB\}$ indica la «probabilidad de $A$ y $B$», es decir, la probabilidad para la unión exterior de $A$ y $B$. Luego,

$$\Pr\{AB\} = \frac{c}{p}.$$

$\Pr\{A \text{ o } B\}$ indica la probabilidad de obtener $A$, o $B$, o ambas $A$ y $B$ *. Luego,

$$\Pr\{A \text{ o } B\} = \frac{a + b + c}{p}.$$

Vemos, realmente, que

$$\Pr\{A\} + \Pr\{B\} = \Pr\{AB\} + \Pr\{A \text{ o } B\}$$

y de aquí se sigue nuestra primera regla fundamental (la regla «o»:)

$$\Pr\{A \text{ o } B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{AB\}. \tag{1}$$

Ahora deseamos definir la probabilidad *condicional* $\Pr\{A/B\}$, en palabras: probabilidad de $A$ *si* $B$ (garantizada $B$, puesta $B$, sobre la condición $B$, sobre la hipótesis $B$, ...). También esta probabilidad se entiende que representa una frecuencia relativa a largo alcance. Saquemos del saco, repetidamente, una bola cada vez, remplazando la bola sacada antes de la extracción anterior, como ha sido descrito en la sección 2 (1). Pero nosotros *sólo tenemos en cuenta las bolas que tienen una* $B$. Si entre las primeras $n$ bolas sacadas hay $m$ bolas que también tienen una $A$, $m/n$ es la frecuencia relativa, que será, aproximadamente,

---

* La palabra «o» tiene dos significados, que no son distinguidos con la claridad necesaria en los modernos idiomas europeos, a pesar de que sí lo eran en latín. Podemos usar «o» en sentido «exclusivo» o «inclusivo». «Tú puedes ir a la playa o al cine» (pero no ambas cosas) es un «o» *exclusivo* (en latín, «aut»). «Tú puedes ir a la playa o tener un paquete de caramelos» es un «o» *inclusivo* si queremos significar «una, la otra o ambas cosas». En los documentos legales o financieros el «o» inclusivo se escribe «y/o» (en latín, «vel»). Cuando escribimos $\Pr\{A \text{ o } B\}$ nosotros queremos significar «o» *inclusivo*.

cuando $n$ es lo bastante grande, igual a $\Pr\{A/B\}$. Esto parece más obvio que

$$\Pr\{A/B\} = \frac{c}{b+c}.$$

En efecto, hay $c$ bolas que tienen $A$ entre las $b + c$ que tienen $B$; aquí el razonamiento de la sección 2 (1) puede ser repetido; desde un cierto punto de vista, podíamos mirar la expresión $\Pr\{A/B\}$ también como una definición. En cualquier caso, podemos encontrar fácilmente, comparando las expresiones de las probabilidades envueltas, que

$$\Pr\{A/B\} = \Pr\{AB\}/\Pr\{B\}.$$

Intercambiando $A$ y $B$ encontramos la segunda regla fundamental (la regla «y»):

$$\Pr\{AB\} = \Pr\{A\}\Pr\{B/A\} = \Pr\{B\}\Pr\{A/B\}. \tag{2}$$

Podemos derivar otras muchas reglas de (1) y (2). Observando que

$$\Pr\{A \text{ o } \bar{A}\} = 1, \qquad \Pr\{A\bar{A}\} = 0,$$

obtenemos de (1), por sustitución de $\bar{A}$ por $B$, que

$$\Pr\{A\} + \Pr\{\bar{A}\} = 1, \tag{3}$$

lo que podríamos ver también directamente, por supuesto. De modo semejante, puesto que

$$\Pr\{AB \text{ o } \bar{A}B\} = \Pr\{B\}, \qquad \Pr\{(AB)\,(\bar{A}B)\} = 0,$$

obtenemos de (1), sustituyendo $AB$ por $A$ y $\bar{A}B$ por $B$ que,

$$\Pr\{B\} = \Pr\{AB\} + \Pr\{\bar{A}B\}. \tag{4}$$

Notamos aquí la siguiente generalización de (2):

$$\Pr\{AB/H\} = \Pr\{A/H\}\Pr\{B/HA\} = \Pr\{B/H\}\Pr\{A/HB\}. \tag{5}$$

También podemos visualizar (5) utilizando el saco y las bolas.

**27.** *Independencia.* Llamamos independientes uno de otro a dos sucesos, si el acaecer (o no acaecer) de uno no tiene influencia en las posibilidades del otro. No obstante, vamos a no considerar por el momento esta definición informal y consideremos las dos definiciones formalessiguientes.

(I) $A$ es llamada *independiente* de $B$ si

$$\Pr\{A/B\} = \Pr\{A/\bar{B}\}.$$

(II) $A$ y $B$ son llamadas *mutuamente independientes* si

$$\Pr\{A/B\} = \Pr\{A/\bar{B}\} = \Pr\{A\}, \qquad \Pr\{B/A\} = \Pr\{B/\bar{A}\} = \Pr\{B\}.$$

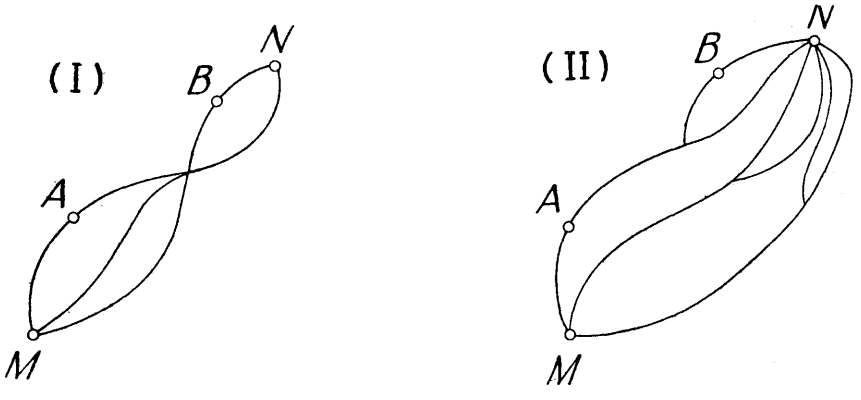



Fig. 14.8. Dos sistemas de carreteras desde la ciudad $M$ a la ciudad $N$, con una diferencia esencial.

Obviamente, si $A$ y $B$ son mutuamente independientes, $A$ es independiente de $B$. Utilizando las reglas del ejemplo 36, probar el teorema: *Si ninguna de las probabilidades* $\Pr\{A\}$, $\Pr\{B\}$, $\Pr\{\bar{A}\}$, $\Pr\{\bar{B}\}$, *desaparece y cualquiera de los dos sucesos* $A$ *y* $B$ *es independiente del otro, éstos son mutuamente independientes.*

**28.** Comparar la sección 3 (5) con el ejemplo 27.

**29.** Un coche que va de la ciudad $M$ a la ciudad $N$ puede pasar por la ciudad $A$ y también por la ciudad $B$. Esto es cierto de ambos sistemas de carreteras, (I) y (II), mostradas en la figura 14.8. Responder a las siguientes preguntas: (*a*), (*b*) y (*c*), suponiendo primero que (I) representa el sistema total de carreteras entre $M$ y $N$, y suponiendo después lo mismo para (II).

(*a*) Supongamos que $A$ denota el hecho de que un coche que va de $M$ a $N$ pasa por la ciudad $A$ y $B$ el que pasa por la ciudad $B$. Supongamos (para los dos sistemas, (I) y (II)) que las tres carreteras que parten de $M$ están igualmente bien frecuentadas (tienen la misma probabilidad) y también que las carreteras que terminan en $N$ (hay dos en (I), seis en (II)) están igualmente bien frecuentadas. Encontrar las probabilidades $\Pr\{A\}$, $\Pr\{A/B\}$, $\Pr\{A/\bar{B}\}$, $\Pr\{B\}$, $\Pr\{B/A\}$, $\Pr\{B/\bar{A}\}$.

(*b*) Encontrar $\Pr\{AB\}$ utilizando la regla (2) del ejemplo 26.

(*c*) Verificar que

$$\Pr\{A\} = \Pr\{B\}\Pr\{A/B\} + \Pr\{\bar{B}\}\Pr\{A/\bar{B}\},$$
$$\Pr\{B\} = \Pr\{A\}\Pr\{B/A\} + \Pr\{\bar{A}\}\Pr\{B/\bar{A}\}.$$

(*d*) ¿Cuál considera la diferencia más importante entre (I) y (II)?

**30.** *Permutaciones a partir de la probabilidad.* Decidir el orden en que los $n$ participantes de un concurso atlético mostrarán su habilidad, el nombre de cada uno está escrito en un pedazo de papel, y luego los $n$ pedazos de papel son sacados de un sombrero, uno tras otro, al azar. ¿Cuál es la probabilidad de que los $n$ nombres aparezcan en orden alfabético?

Presentamos dos soluciones y sacamos una conclusión, comparándolas.

(1) Llamamos $E_1$ el hecho de que el papel sacado primero sea también el primero alfabéticamente, $E_2$ el hecho de que el papel sacado en segundo lugar sea también alfabéticamente el segundo, y así sucesivamente. La probabilidad deseada es

$$\Pr\{E_1 E_2 E_3 \ldots E_n\} =$$
$$= \Pr\{E_1\}\Pr\{E_2/E_1\}\Pr\{E_3/E_1 E_2\} \ldots \Pr\{E_n/E_1 \ldots E_{n-1}\}$$
$$= \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} \cdots \frac{1}{1}.$$

De hecho, obtenemos la primera transformación aplicando las reglas (2) y (5) del ejemplo 26, y la segunda transformación observando que hay $n$ casos posibles para $E_1$, $n-1$ para $E_2$ tras $E_1$, $n-2$ para $E_3$ tras $E_1$ y $E_2$, y así sucesivamente, mientras para cada uno de estos hechos hay un solo caso favorable.

(2) Llamemos $P_n$ el número de todas las posibles ordenaciones (permutaciones, colocaciones lineales, ...) de $n$ objetos distintos. Los $n$ nombres pueden salir del sombrero de $n$ modos, ninguno de estos $P_n$ casos posibles parece más privilegiado que los otros y entre estos $P_n$ casos sólo uno es favorable (el orden alfabético). Por tanto, la probabilidad deseada es $1/P_n$.

(3) Los resultados derivados en (1) y (2) deben ser iguales. Igualándolos evaluamos $P_n$:

$$P_n = 1 \cdot 2 \cdot 3 \ldots n = n!$$

**31.** *Combinaciones a partir de la probabilidad.* La señora Smith compró $n$ huevos, sin comprobar que $r$ de éstos estaban podridos.

Necesitaba $r$ huevos y eligió al azar estos $r$ huevos entre los $n$ que tenía. ¿Cuál es la probabilidad de que los $r$ huevos elegidos sean los podridos?

Como en el ejemplo 30, presentamos dos soluciones y sacamos una conclusión, comparándolas.

(1) Llamemos $E_1$ el hecho de que el primer huevo abierto por la señora Smith esté podrido, $E_2$ el hecho de que el segundo huevo esté podrido, y así sucesivamente. La probabilidad deseada es

$$\begin{aligned}&\Pr\{E_1 E_2 E_3 \dots E_r\}\\ &\quad = \Pr\{E_1\}\ \Pr\{E_2/E_1\}\ \Pr\{E_3/E_1E_2\}\ \dots\ \Pr\{E_r/E_1 \dots E_{r-1}\}\\ &\quad = \frac{r}{n}\cdot\frac{r-1}{n-1}\cdot\frac{r-2}{n-2}\cdots\frac{1}{n-r+1}\ .\end{aligned}$$

La primera transformación es obtenida por la reglas (2) y (5) del ejemplo 26, la segunda de la consideración de los casos posibles y favorables para $E_1$, para $E_2$ tras $E_1$, y así sucesivamente.

(2) Tenemos un conjunto de $n$ objetos distintos. Cualesquiera $r$ objetos elegidos entre estos $n$ objetos forman un subconjunto de tamaño $r$ del conjunto dado de tamaño $n$: llamamos $C_r^n$ al número de todos estos subconjuntos. (Generalmente, se llama $C_r^n$ al número de «combinaciones» de $r$ cosas seleccionadas de entre $n$ cosas.) En el caso de los huevos de la señora Smith hay $C_r^n$ casos posibles, ninguno de ellos con más privilegios que los otros, y entre estos $C_r^n$ casos, sólo uno es favorable (si conseguir huevos podridos lo es). De aquí que la deseada probabilidad sea $1/C_r^n$.

(3) Comparando (1) y (2) evaluamos $C_r^n$:

$$C_r^n = \frac{n(n-1)\ \dots\ (n-r+1)}{1\cdot 2\quad\dots\quad r} = \frac{n!}{r!\,(n-r)!} = \binom{n}{r}.$$

**32.** *La elección de una conjetura estadística rival: un ejemplo.* Una persona retiró 875 dólares de su cuenta de ahorro en cierta fecha, y otra persona recibió 875 dólares dos días después. La coincidencia de estas dos cantidades, una retirada y otra recibida, puede ser considerada como una evidencia circunstancial, como una indicación de que se ha cometido un delito; cf. ejemplo 13.6. Si el jurado encuentra difícil creer que esta coincidencia se debe al mero azar, puede dar por resultado una declaración de culpabilidad. De aquí el problema siguiente: ¿Cuál es la probabilidad de una coincidencia semejante? Cuanto menor es la probabilidad más difícil es atribuir la coincidencia al azar y más fuerte la evidencia en contra de los demandados.

Sin embargo, no podemos calcular una probabilidad numérica sin asumir alguna hipótesis estadística definida. ¿Qué hipótesis asumiremos? En un caso serio deberemos asumir serios pensamientos. Examinemos unas cuantas posibilidades.

(1) Como el número 875 tiene tres dígitos podemos considerar los enteros positivos, que no tienen más de tres dígitos, como admisibles, y considerarlos a todos admisibles por igual. La probabilidad de que dos de tales enteros, elegidos al azar, independientemente uno de otro, coincidan es obviamente 1/999. Esta probabilidad es bastante pequeña; pero ¿es razonable la suposición que subyace en su cálculo?

(2) Como 875 tiene menos de cinco dígitos tenemos que considerar todos los enteros como menos de cinco dígitos como igualmente admisibles. Esto conduce a la probabilidad 1/9999 de la coincidencia. Esta probabilidad es muy pequeña realmente, pero nuestra suposición viene traída por los pelos, es incluso frívola.

(3) Si el caso es importante, el tribunal puede ordenar la inspección de los libros del Banco o convocar a uno de sus oficiales competentes para atestiguar. Y así, hemos averiguado que inmediatamente antes de retirar los 875 dólares se había depositado la cantidad de 2581,48 dólares en la misma cuenta. Con esta importante información podemos considerar como casos posibles e igualmente admisibles las sumas 1, 2, 3, ... 2581 que podían haber sido retiradas de la cuenta. Sólo se dio uno de estos casos, el 875, y así nos vemos conducidos a la probabilidad 1/2581 de la coincidencia. Esta probabilidad es pequeña, pero nuestra suposición parece razonablemente realista.

(4) Podríamos haber considerado no sólo los retiros en dólares, sino también en dólares y centavos, como 875,31 dólares. Si consideramos todos estos casos igualmente admisibles, la probabilidad de una coincidencia llega a ser 1/258148. Esta probabilidad es muy pequeña, pero nuestra suposición parece menos realista: los retiros en dólares y centavos, como el de 875,31 dólares, son más normales en una cuenta corriente que en una de ahorro.

(5) Por el contrario, se podría argüir que las cantidades retiradas de una cuenta de ahorro son normalmente «redondas», divisibles por 100, 50 ó 25. Ahora bien, 875 es divisible por 25. Si consideramos admisibles sólo los múltiplos de 25, y todos ellos igualmente admisibles, la probabilidad en cuestión llega a ser de 1/103.

Por supuesto, podríamos imaginar otras y más complicadas maneras de calcular la probabilidad, pero no insistiremos en un ejemplo tan claro. El ejemplo ha cumplido su propósito si el lector puede comprender ahora los dos puntos siguientes.

(*a*) Aunque algunas de las cinco suposiciones examinadas pueden ser más aceptables que otras, ninguna es claramente superior a las demás, y hay poca esperanza de encontrar una suposición que sea aceptable en todos los aspectos, y pueda ser considerada la mejor.

(*b*) Cada una de las cinco suposiciones consideradas admite poca probabilidad para la coincidencia observada, y así, en el conjunto, nuestra consideración se mantiene en el punto de vista del sentido común: «es difícil creer que esta coincidencia se deba al mero azar».

**33.** *La elección de una conjetura estadística rival: observaciones generales.* Intentemos aprender algo más general del ejemplo particular considerado (ejemplo 32). Volvamos a considerar la situación general examinada en la sección 14.9 (7). Un suceso $S$ ha ocurrido y ha sido observado. Hay dos conjeturas rivales, que se refieren a este suceso: una «física», $F$, y una hipótesis estadística, $H$. Si aceptamos la conjetura física $F$, $S$ se explica fácilmente. Si aceptamos la hipótesis estadística $H$, podremos calcular la probabilidad $p$ del acaecer del suceso $S$. Si $p$ es «pequeña», podemos vernos inducidos a rechazar la hipótesis estadística $H$. En cualquier caso, la pequeñez de $p$ debilita nuestra confianza en $H$ y, por tanto, fortalece de algún modo nuestra confianza en la conjetura rival $F$.

Sin embargo, el ejemplo 32 nos hace conscientes de que la calidad de la hipótesis estadística $H$ juega un papel en el razonamiento descrito. La hipótesis estadística $H$ puede parecernos innatural, inapropiada, traída por los pelos, frívola, gratuita, *indigna de confianza* desde el principio. O, por el contrario, $H$ puede parecer natural, apropiada, realista, razonable, *digna de confianza.*

Ahora bien, $p$, la probabilidad del suceso $S$, calculado sobre la base de la hipótesis $H$, puede ser tan pequeña que rechacemos $H$: una rival de $F$ desaparece de la competición. Esto aumenta las perspectivas de $F$ —pero puede aumentarlas mucho o poco: esto depende de la calidad de la rival desaparecida—. Si la hipótesis estadística $H$ se nos aparece, desde el principio, como apropiada, digna de confianza, $H$ fue una rival peligrosa y su caída fortalece a $F$ considerablemente. Si, no obstante, $H$ nos pareció inapropiada y de poca confianza desde el principio, $H$ fue una rival débil; su caída no nos sorprende y fortalece a $F$ muy poco.

Dada una hipótesis estadística clara $H$, la probabilidad $p$ del suceso $S$ está determinada, y el estadístico puede calcularla. Pero el cliente del estadístico, que puede ser un biólogo, o un hombre de negocios, o cualquier otro no estadístico, ha de decidir lo que este valor numérico de $p$ pequeña significa en su caso. Ha de decidir el grado de pequeñez de $p$,

que basta para rechazar o debilitar la hipótesis estadística *H*. Pero el cliente no está normalmente interesado, de modo directo, en dicha hipótesis estadística; a éste le concierne, de modo primario, la conjetura «física», rival, *F*. Y ha de decidir lo que pesa el rechazo o la debilitación de *H* en el fortalecimiento de *F*. Esta última decisión no puede depender sólo del valor numérico de *p*, como es natural: depende, ciertamente, de la evaluación de *H*.

Me temo que el cliente del estadístico, que desee hacer uso del valor numérico *p*, sin comprobar la importancia de la hipótesis estadística *H* para su problema, se engaña a sí mismo. Podrá darse escasamente cuenta de la importancia de *H* si no comprueba que su conjetura física *F* puede ser confrontada con hipótesis estadísticas diferentes de *H*. Cf. ejemplo 15.5.

# El cálculo de probabilidad y la lógica del razonamiento plausible

*Es difícil estimar la probabilidad de los resultados de la inducción.*—LAPLACE *.

*Sabemos que la probabilidad de una inducción bien establecida es grande, pero cuando se nos pide que fijemos su grado no podemos hacerlo. El sentido común nos dice que unos argumentos inductivos son más fuertes que otros y que algunos lo son mucho. Sin embargo, no podemos expresar cuánta más fuerza tienen unos que otros ni su grado de fortaleza.*—JOHN MAYNARD KEYNES **.

## 1. REGLAS DE RAZONAMIENTO PLAUSIBLE

En los tres capítulos anteriores hemos recogido patrones de razonamiento plausible. ¿Constituyen estos «patrones» «reglas» para esta clase de razonamiento? ¿Hasta qué punto y de qué manera son obligatorios e imperativos? Corremos el peligro de perdernos en explicaciones puramente verbales. Por tanto, prefiero considerar la cuestión de modo más concreto, incluso un poco personalmente.

(1) Recuerdo una conversación sobre invención y razonamiento plausible, ocurrida hace bastante tiempo. Hablaba con un amigo mucho mayor que yo, el cual tenía tras de sí un distinguido expediente de descubrimientos, inventos y afortunado trabajo profesional. Sin duda, cuando hablaba de razonamiento plausible e invención sabía de lo que estaba hablando. Sostenía con ardor y fuerza de convicción poco usuales que tanto el uno como la otra no tenían reglas. Presentimientos y conjeturas dependen de la intuición y la experiencia, decía: no hay reglas, no puede haberlas, ni las habrá, y, si las hubiera, serían inútiles, en cualquier caso. Yo sostenía lo contrario-una conversación carece de interés si

* *Essai philosophique sur les probabilités.* Véase *Oeuvres complètes de Laplace*, vol. 7, página. CXXXIX.

** *A treitise on probability*, pág. 259.

no hay diferencia de opiniones—, aun cuando sentía la fuerza de su posición. Mi amigo era un cirujano. Una decisión equivocada suya podía costar una vida y, a veces, cuando un paciente empieza a sangrar o a ahogarse, hay que decidir en un segundo. Comprendí que quienes tienen que decidir tan rápidamente en asuntos de responsabilidad no usan reglas. El tiempo es demasiado escaso para aplicar bien una regla, y una idea general podría llevarles a un error; lo que se necesita es concentrarse intensamente en la situación que se tiene delante. Por eso la gente llega a desconfiar de las «reglas» y confía en su «intuición», o en su «experiencia», o en ambas a la vez.

En el caso de mi amigo quizá había algo más. Siendo persona un poco dominante, odiaba renunciar a su poder. Tal vez sentía que el reconocimiento de una regla era como delegar una parte de su autoridad en una máquina, y, por eso, estaba en contra de ellas.

Tomemos, pues, nota de que gente inteligente puede desconfiar de las reglas de razonamiento.

(2) Dos personas situadas ante la misma evidencia pueden juzgarla de modo muy diferente. Dos miembros de un tribunal que trabajan en el mismo expediente pueden estar en desacuerdo; uno piensa que la evidencia presentada es una prueba suficiente contra el demandado y el otro piensa que no lo es. Tal desacuerdo puede tener mil fundamentos distintos: la gente puede moverse en direcciones opuestas por temores, esperanzas, prejuicios y simpatías, o por diferencias personales. Quizá uno de los miembros del tribunal es un estúpido, mientras que el otro es inteligente, o uno duerme durante las vistas y el otro escucha atentamente. Pero las diferencias personales que subyacen al desacuerdo pueden ser más sutiles. Quizá ambos miembros del tribunal son honrados y sin prejuicios, ambos siguen las vistas de los casos con atención, y son los dos inteligentes, pero de modo diferente. El primero puede ser mejor observador que el otro del aspecto externo. Observa la expresión del rostro de los testigos, los tics del demandado, y cuándo una respuesta es dada de modo inseguro; los movimientos rápidos de los ojos y el gesto parco de las manos le impresionan. El segundo puede tener menos habilidad para observar las expresiones de los rostros, pero es mejor juez de las relaciones sociales: comprende mejor el medio y las circunstancias de la gente complicada en un caso. Viendo las mismas cosas con ojos diferentes, honradamente y no sin inteligencia, los dos miembros del tribunal llegan a conclusiones opuestas.

Tomemos nota de nuevo sin desdeñar las cuestiones obvias: dos personas pueden estar en desacuerdo ante la misma evidencia.

(3) Mi amigo y yo estamos interesados en la conjetura $A$. (Este

amigo es un matemático, y $A$ es una conjetura matemática.) Los dos sabemos que $A$ implica a $B$. Y ahora encontramos que $B$, consecuencia de $A$, es verdadera. Estamos de acuerdo que esta verificación de su consecuencia $B$ es una evidencia en favor de la conjetura $A$, pero estamos en desacuerdo respecto al valor o peso de esta evidencia. Uno de nosotros afirma que esta verificación añade muy poco al crédito de $A$ y el otro afirma que añade mucho.

Este desacuerdo sería comprensible si no estuviéramos igualmente familiarizados con el tema y uno de nosotros conociese muchas más consecuencias primeramente verificadas que el otro. Pero esto no es así. Conocemos aproximadamente las mismas consecuencias de $A$ verificadas en el pasado. Estamos incluso de acuerdo en que hay poca analogía entre $B$, que acaba de ser verificada, y aquellas consecuencias verificadas primeramente. Coincidiendo en que esta circunstancia hace más fuerte la evidencia de $A$. Pero uno de nosotros dice: «sólo un poco más fuerte», y el otro: «mucho más fuerte», y aquí está el desacuerdo.

Ambos sospechábamos, hacía poco, que $B$ era falsa y fue una sorpresa descubrir lo contrario. En efecto, desde el punto de partida de una suposición más que natural (o hipótesis estadística) $B$ parece bastante improbable. Los dos percibimos que esta circunstancia hace más fuerte la evidencia de $A$. Sin embargo, uno dice: «sólo un poco más fuerte», y el otro: «mucho más fuerte», y continuamos en desacuerdo.

Siendo los dos perfectamente honestos, creo, y no siendo nuestro desacuerdo una mera cuestión de temperamentos, diferimos a causa de nuestras posiciones. Aunque tenemos aproximadamente el mismo adiestramiento científico, nos hemos desarrollado en diferentes direcciones. Su trabajo le lleva a desconfiar de la hipótesis $A$. Quizá espera poder refutar un día esta conjetura. En cuanto a mí, no me atrevo a esperar que un día podré probarla. Sin embargo, debo confesar que me gustaría hacerlo. En realidad, ésa es mi ambición, pero no deseo caer en la ilusión de que podré hacerlo. Esta esperanza apenas confesada puede influir en mi juicio, y en mi evaluación del peso de la evidencia. No obstante, aún puedo tener otras razones, quizá más oscuras, escasamente formuladas e inarticuladas. Y mi amigo puede tener también algunas razones, que no se confiesa a sí mismo. De cualquier modo, tales diferencias en nuestra posición pueden explicar la situación: diferimos en lo referente a la fuerza de la evidencia, aunque estamos de acuerdo en todos los puntos claramente reconocidos, que influyen en la fuerza de la evidencia de un juicio impersonal según normas universalmente aceptadas.

Observemos, pues, que dos personas puestas ante la misma eviden-

cia, y aplicando los mismos patrones de inferencia plausible, pueden diferir de opinión.

(4) Hemos intentado ver el razonamiento plausible concretamente, en funciones, en la conducta de la gente que se enfrenta a problemas concretos. Tenemos ahora, según espero, una idea algo más clara de la manera en que nuestros patrones «obligan», y hasta qué punto pueden considerarse como «reglas».

Pero quedan por examinar otros planteamientos. La lógica formal y el cálculo de probabilidades tienen reglas claras y estrictas que parecen estar relacionadas con nuestros patrones. ¿Cuál es la naturaleza de esta relación? Esta es la cuestión que discutiremos en las próximas secciones.

## 2. Un aspecto del razonamiento demostrativo

En esta etapa puede ser útil una comparación del razonamiento plausible con el razonamiento demostrativo. Pero, naturalmente, el nivel a que ha llegado el razonamiento plausible no soporta la comparación con la etapa altamente desarrollada a que la teoría del razonamiento demostrativo ha llegado en esta época, después de un desarrollo de más de dos mil años, de los que los últimos cincuenta han sido especialmente importantes en este sentido. Será mejor que hagamos la comparación con un aspecto primitivo del razonamiento plausible. Pongámonos, más o menos, en la posición de un contemporáneo de Aristóteles.

Este observó que el razonamiento se produce conforme a ciertos patrones; imagino que percibió tales patrones en los argumentos políticos, filosóficos, legales e incluso en los de cada día, y que, reconociéndolos mientras se producían, los extrajo y formuló. Estos patrones son los silogismos. Los ejemplos con los que Aristóteles ilustra sus silogismos parecen aducir la idea de que descubrió sus silogismos por una especie de inducción —y ¿cómo podría haber sido de otra manera?—. En cualquier caso, la idea de que los silogismos pueden haberse descubierto inductivamente los acerca algo a nuestros patrones de razonamiento plausible.

En lugar del silogismo «subsuntivo», tan querido de Aristóteles y aún más de sus seguidores escolásticos, nosotros consideraremos el *modus tollens* del silogismo «hipotético», que hemos considerado ya en la sección 12.1:

$A$ implica $B$
$B$ falsa
---
$A$ falsa.

Incluso desde un punto de vista bastante primitivo podemos observar varios rasgos notables en este patrón de razonamiento: es *impersonal, universal, autosuficiente* y *definitivo.*

(1) Al usar la palabra *impersonal* hacemos hincapié en que la validez del razonamiento no depende de la personalidad del razonador, cualquiera que sea su humor, su gusto, su clase, su credo o su color.

(2) Con la palabra *universal* hacemos hincapié en que los enunciados considerados (denotados por $A$ y $B$) no necesitan pertenecer a este o aquel campo del conocimiento, ya sea las matemáticas o la física, ya sea la lógica o el derecho, sino que puede pertenecer a uno de estos campos o a cualquier otro. Pueden referirse a cualquier objeto del pensamiento humano la suficientemente claro: la conclusión silogística se aplica a toda clase de objetos semejantes.

(3) Para comprender el próximo punto debemos darnos cuenta de que nuestro conocimiento y nuestras creencias razonables pueden ser cambiadas por una información nueva. Sin embargo, hay algo invariable en el silogismo considerado. Una vez aceptadas las premisas no podemos dejar de aceptar la conclusión. Algún tiempo más tarde podemos recibir nueva información sobre los temas implicados en nuestro silogismo. Si, no obstante, esta información no varía la aceptación de las premisas no puede tampoco variar razonablemente la aceptación de la conclusión. La inferencia de un silogismo demostrativo no necesita nada exterior, es independiente de todo lo que no esté explícitamente mencionado en las premisas. En este sentido, el silogismo es *autosuficiente;* nada se necesita, aparte de las premisas, para sacar la conclusión y nada puede invalidarla si las premisas son sólidas.

Esta «autosuficiencia» o «autarquía» del silogismo es, quizá, su rasgo más notable. Citemos al mismo Aristóteles: «Un silogismo es un discurso en el que, habiendo establecido ciertas cosas, otras distintas se siguen necesariamente de aquéllas. Por esto último entiendo que éstas producen la consecuencia, y por esto otro, que no se requiere ningún nuevo término para hacer la consecuencia necesaria».

(4) Si las premisas son incotestablemente ciertas, podemos «separar» la conclusión del silogismo. Es decir, si estamos seguros de que tanto «A implica $B$» como «$B$ es falsa» son ciertas, podemos olvidarnos de estas premisas y quedarnos con la conclusión «$A$ es falsa» como posesión mental *definitiva.*

Hemos examinado sólo una de las varias clases de silogismos, pero se trata precisamente de uno típico: también los otros silogismos son impersonales, universales, autosuficientes y definitivos. Y estos rasgos simbolizan el carácter general del razonamiento demostrativo.

### 3. Un aspecto similar del razonamiento plausible

Comparemos el patrón de razonamiento demostrativo (el *modus tollens*) examinado en la sección anterior con el patrón de razonamiento plausible presentado en la sección 12.1:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ cierta}\end{array}}{A \text{ más digna de crédito.}}$$

Entre estos dos patrones, el «demostrativo» y el «plausible», hay una cierta similaridad externa. (El patrón demostrativo es tradicional, mientras el otro se ha adoptado después.) Pero, comparémolos de modo más completo.

Ambos patrones tienen la misma primera premisa

$$A \text{ implica } B.$$

Las segundas premisas

$$B \text{ falsa} \qquad\qquad B \text{ cierta}$$

son opuestas, aunque igualmente claras y definitivas; están en el mismo nivel lógico. Sin embargo, hay una gran diferencia entre las dos conclusiones

$$A \text{ falsa} \qquad\qquad A \text{ más digna de crédito.}$$

Estas conclusiones están en diferentes niveles lógicos. La conclusión del patrón demostrativo está al mismo nivel que las premisas, mientras la conclusión del patrón de razonamiento plausible es de naturaleza diferente, menos fuerte, mucho peor expresada.

La conclusión plausible es comparable a una fuerza con *dirección y magnitud.* La conclusión nos empuja en una cierta dirección: *A* se hace *más* digna de crédito. Pero esa misma conclusión tiene también cierta fuerza: *A* puede alcanzar *mucho más* crédito o *sólo un poco* más de crédito. La conclusión no está completamente expresada ni apoyada por las premisas. *La dirección está expresada e implicada en las premisas, la fortaleza no.* Para cualquier persona razonable las premisas implican que *A* se hace más digna de crédito (no menos, desde luego), pero mi amigo y yo podemos diferir en *cuánto* más digna de crédito. *La dirección es impersonal, la fortaleza puede ser personal.* Mi amigo y yo podemos estar

honestamente en desacuerdo sobre el peso de la conclusión, puesto que nuestro temperamento, nuestras posiciones y nuestras razones sin formular pueden ser diferentes. Pero la fuerza de la conclusión es importante. Si dos jueces juzgan de modo distinto la fuerza de una conclusión, uno puede estar por la absolución y el otro en contra. Si dos científicos juzgan de modo diferente la fuerza de una conclusión, uno puede estar en pro de cierto experimento y el otro en contra.

La conclusión de nuestro patrón de inferencia plausible parece *unilateral* cuando le comparamos con las creencias y los actos reales de las personas que razonan: expresa un solo aspecto y descuida los otros. Si comprobamos esto, la naturaleza del razonamiento plausible puede parecer menos desconcertante y alusiva. En cualquier caso, estaremos mejor preparados para comparar los patrones, el demostrativo y el plausible, punto por punto. Cada una de las siguientes subsecciones se corresponden con las subsecciones numeradas de la anterior sección 2

(1) Cuando razonamos de acuerdo con nuestro patrón de inferencia plausible nos ajustamos a un principio: la verificación de una consecuencia fortalece la conjetura. Este principio parece ser reconocido independientemente de las diferencias personales y de las idiosincrasias. Por ello nuestro patrón resulta *impersonal.*

No obstante, pagamos un precio por tal «impersonalidad». Nuestros patrones logran ser impersonales, porque son unilaterales, restringidos a un aspecto de la inferencia plausible. En cuanto hacemos la pregunta «¿*Cuánto* resulta fortalecida la conjetura por la verificación de esta consecuencia?» abrimos la puerta a las diferencias personales.

(2) Nos hemos tomado el trabajo de mostrar en anteriores capítulos, por medio de ejemplos, que al tratar con conjeturas matemáticas seguimos de modo natural nuestro patrón de inferencia plausible. El principio subyacente es generalmente reconocido en ciencias naturales, admitido implícitamente por los tribunales de la ley y en las cuestiones de la vida diaria. La verificación de una consecuencia es considerada como evidencia razonable de una conjetura en cualquier dominio. Por ello nuestro patrón resulta *universal.*

No obstante, pagamos un precio por tal «universalidad». Nuestro patrón logra ser universal porque es unilateral, restringido a un solo aspecto de la inferencia plausible. La universalidad se empaña cuando preguntamos: «¿Cuál es el peso de una evidencia semejante?». Al objeto de juzgar el peso de la evidencia debemos estar familiarizados con la materia. Para juzgar su peso con seguridad hemos de ser expertos en ella. No obstante, no podemos estar familiarizados con todas las materias y menos aún ser expertos en todas ellas. Así, cada uno de nos-

otros observará pronto que existen límites prácticos para la universalidad de la inferencia plausible.

(3) Hasta donde se manifiesta, la conclusión plausible está apoyada por las premisas. Sobre la base de la evidencia proporcionada por las premisas es razonable poner más confianza en $A$. En alguna fecha posterior, sin embargo, podemos recibir nueva información que, sin cambiar nuestra confianza en las premisas, pueda cambiar nuestra opinión respecto de $A$: podemos encontrar a $A$ menos digna de crédito o incluso podemos probar que $A$ es falsa.

Esto no constituye una objeción contra el patrón de razonamiento. Hasta donde llega la evidencia expresada en las premisas la conclusión resulta justificada. Un veredicto del jurado puede condenar al inocente o absolver al delincuente. Sin embargo, la injusticia del veredicto puede tener justificación, puesto que, basándose en la evidencia disponible, no era posible dar un veredicto mejor. Tal es la naturaleza de la inferencia plausible, y así, nuestro patrón de razonamiento plausible puede ser denominado *autosuficiente*.

Esta clase de autosuficiencia o «autarquía» no significa durabilidad. Por otra parte, el peso de la evidencia (que no se menciona en la conclusión de nuestro [unilateral] patrón, pero que es importante, no obstante) depende de cosas que no están mencionadas en las premisas. La fuerza de la conclusión (no su dirección) requiere cosas externas a las premisas.

(4) No podemos «separar» la conclusión de nuestro patrón de razonamiento plausible. «$A$ se ha hecho más digna de crédito» no tiene sentido sin referencia a las premisas que explican las circunstancias de las que surgió. Referida a las premisas, la conclusión plausible goza de claro sentido y resulta perfectamente razonable, pero puede disminuir el valor con el tiempo, aunque las premisas permanezcan intactas. La conclusión plausible puede ser muy valiosa en el momento en que se produce, pero nuevos conocimientos pueden depreciarla: su importancia es sólo momentánea, transitoria, efímera, *provisional*.

En resumen: nuestro patrón de razonamiento plausible es unilateral y deja un amplio margen para el desacuerdo en cosas de importancia. No obstante, al precio de esta unilateralidad, consigue ser impersonal y universal, incluso autosuficiente, hasta cierto punto. Aun así, no puede escapar de ser meramente provisional.

Sería locura lamentar que en varios aspectos nuestro patrón de razonamiento plausible no llegue a la perfección del razonamiento demostrativo. Por el contrario, sentiríamos poca satisfacción al lograr aclarar algo una diferencia que podíamos haber sospechado desde el principio.

Desde el comienzo quedó claro que las dos clases de razonamiento tienen tareas diferentes. También desde el principio parecían muy diferentes: el razonamiento demostrativo, como preciso, final, «automático»; y el razonamiento plausible, como vago, provisional y específicamente «humano». Ahora podemos ver la diferencia un poco más distintamente. En oposición a la inferencia demostrativa, la inferencia plausible deja indeterminada una cuestión de gran importancia: la «fuerza» o el «peso» de la conclusión. Este peso puede depender no sólo de bases claras como las expresadas en las premisas, sino también de bases que no están expresas ni aclaradas y pertenecen a la posición de la persona que saca la conclusión. Una persona toma una posición, una máquina no hace esto. En realidad, se puede construir una máquina que saque conclusiones demostrativas por nosotros, pero creo que no se podrá nunca construir una máquina que obtenga inferencias plausibles.

## 4. Un aspecto del cálculo de probabilidad. Dificultades

Un paso muy importante en la construcción de una teoría física es su *formulación en términos matemáticos*. Hemos llegado a un punto en nuestra investigación en que debemos emprender semejante paso; ahora debíamos formular nuestros puntos de vista sobre el razonamiento plausible en términos matemáticos.

Ningún intento de formular una teoría de razonamiento plausible puede desconocer un hecho histórico: el cálculo de probabilidad fue considerado por Laplace y muchos otros científicos eminentes como la expresión adecuada de las reglas de inferencia plausible. Hay motivos para esta opinión, aunque también existen objeciones en contra. Empecemos mejor con las dificultades.

Deseamos utilizar el cálculo de probabilidad para hacer más precisos nuestros puntos de vista sobre el razonamiento plausible. Sin embargo, podemos tener algunas dudas sobre semejante procedimiento, pues hemos visto en el capítulo anterior que el cálculo de probabilidad es una teoría (bastante aceptable) de los fenómenos azarosos de masas. ¿Cómo podría ser el cálculo de probabilidad la teoría de los fenómenos de masas y la lógica de la inferencia plausible?

Esto no es una objeción fuerte; no hay dificultad real. El cálculo de probabilidad *podría* ser ambas cosas, *podría* tener dos interpretaciones. De hecho, una teoría matemática puede tener varias interpretaciones diferentes. La misma ecuación diferencial (la ecuación de Laplace) describe la corriente continua y no rotativa de un fluido incomprensible

y no viscoso, y también la distribución de fuerzas en un campo electrostático. La misma ecuación describe también la corriente continua del calor, la de la electricidad, la disolución de la sal en agua bajo condiciones apropiadas, y todavía algunos otros fenómenos. Y de aquí que no debamos excluir *a priori* que una teoría matemática puede servir a dos propósitos. Quizá podamos utilizar el cálculo de probabilidad tanto para describir los fenómenos azarosos de masas como para sistematizar nuestras reglas de inferencia plausible.

Es importante, no obstante, distinguir con claridad entre estas dos interpretaciones. Así, podemos usar el símbolo $\Pr\{A\}$ (ver ejemplo 14.26) para ambas interpretaciones, aunque con ciertas precauciones, comprendiendo claramente ambos significados del símbolo y ver las diferencias entre ambos significados.

En el capítulo anterior sobre los fenómenos azarosos de masas consideramos algunos tipos de sucesos $A$, como el nacimiento de un niño, la caída de las gotas de lluvia en un lugar específico, el sacar cierto número de puntos con un dado, y así sucesivamente. Nosotros utilizamos el símbolo $\Pr\{A\}$ para denotar la probabilidad del suceso $A$, es decir, el valor teórico de la frecuencia relativa de largo alcance del suceso $A$.

No obstante, en el capítulo presente tenemos que tratar con el razonamiento plausible. Consideraremos una conjetura $A$, y nos veremos enfrentados con la confiabilidad de esta conjetura $A$, la fuerza de la evidencia en favor de $A$, nuestra confianza en $A$, el grado de creencia que daremos a $A$; en una palabra, la *credibilidad de la conjetura A*. Y nosotros tomaremos el símbolo $\Pr\{A\}$ para denotar la credibilidad de $A$.

Así, en el presente capítulo usaremos el símbolo $\Pr\{A\}$ en su segundo significado, a menos que se indique explícitamente lo contrario. Tal uso del símbolo no es objetable, pero tendremos que discutir cuidadosamente el concepto de credibilidad si no deseamos exponernos a graves objeciones.

Primero, debemos evitar una ambigüedad. El símbolo $\Pr\{A\}$ representará la credibilidad de $A$, o la fuerza de la evidencia de la conjetura $A$. Esta evidencia es fuerte si es convincente, y es convincente si convence a alguien. Pero no decimos a quién tiene que convencer: ¿a usted, a mí, al señor Smith o a la señora Jones, o a quién? La fuerza de la evidencia podría también concebirse de modo *impersonal*. Si la concebimos de esta manera, el grado de creencia que usted o yo o cualquiera otra persona tenga en una conjetura propuesta carece de importancia; lo que importa es el *grado de creencia razonable* que cualquiera de nos-

otros *tendría*. Todavía no hemos decidido, y hemos de decidir, en qué sentido exacto hemos de usar el término «credibilidad de $A$» y el símbolo correspondiente $\mathrm{Pr}\{A\}$.

Hay otra dificultad. Las magnitudes consideradas por los físicos, tales como «masa», «carga eléctrica» o «velocidad de reacción», tienen una definición *operacional;* el físico conoce exactamente qué operaciones tiene que realizar si desea hallar la magnitud de una carga eléctrica, por ejemplo. La definición de «frecuencia relativa de largo alcance», aunque de alguna manera menos precisa que la de una carga eléctrica, es aún operacional; esto sugiere operaciones definidas que pueden ser emprendidas para obtener un valor numérico aproximado de tal frecuencia. El problema que se plantea con el concepto de «credibilidad de una conjetura» es que no conocemos ninguna definición operacional de él. ¿Cuál es la credibilidad de la conjetura de que el señor Jones es infiel a su esposa?

Esta credibilidad puede tener en este momento un valor definido en la mente de la señora Jones (confiamos que un valor insignificante), pero no sabemos cómo determinarlo numéricamente. ¿Cuál es la credibilidad de la ley de gravitación universal juzgada sobre la base de las observaciones que se encuentran en la primera edición de los *Principia* de Newton? Esta pregunta podría tener gran interés para algunos de nosotros. (No para la señora Jones, quizá, sino para Laplace o Keynes, si estuviesen aún vivos —ver las citas puestas al principio de este capítulo—.) Pero nadie osaría proponer un valor numérico definitivo para una credibilidad semejante.

Tenemos aún que dar una interpretación adecuada del término «credibilidad de la conjetura $A$» y el símbolo correspondiente $\mathrm{Pr}\{A\}$. Esta interpretación debe ser tal que la dificultad de una definición operacional no interfiera con ella. Por lo demás, y esto es lo principal, esta interpretación nos capacitaría para considerar las reglas del razonamiento plausible de modo sistemático y realista.

## 5. Un aspecto del cálculo de probabilidad. Un intento

Le acaban de presentar a usted a don Fulano de Tal y tiene muy pocas palabras que decirle. Son completamente extraños el uno al otro y la conversación será precavida. Sin embargo, usted no podrá dejar de hacer algunas afirmaciones, tales como «mañana lloverá», «el próximo partido lo ganarán los Blues», «la corporación Esto-y-lo-otro pagará mayores dividendos el próximo año», «la señora Mengana, de cuyo di-

vorcio habla toda la ciudad, fue infiel», «la polio es causada por un virus», o cualesquiera otras afirmaciones $A$, $B$, $C$, $D$, $E$, ... Don Fulano de Tal dará a la afirmación $A$ un grado determinado de creencia $\Pr\{A\}$. Si usted es muy listo, después de estar algún tiempo con don Fulano de Tal podrá darse cuenta de si $\Pr\{A\}$ es alta o baja. No obstante, por muy listo que sea, no creo que sea capaz de dar un valor numérico determinado a $\Pr\{A\}$, la credibilidad del enunciado $A$, por los ojos de este señor. (Aunque podría ser interesante: los valores de $\Pr\{A\}$, $\Pr\{B\}$, $\Pr\{C\}$, ... podrían caracterizar agudamente la personalidad de don Fulano de Tal.)

Seamos realistas y reconozcamos la imposibilidad de una tarea, que está obviamente más allá de nuestros medios: consideremos $\Pr\{A\}$, la credibilidad de la conjetura $A$, por los ojos de don Fulano de Tal, como una fracción determinada

$$0 < \Pr\{A\} < 1.$$

El valor numérico, que, no obstante, *desconocemos*. Y tratemos de modo semejante a $\Pr\{B\}$, $\Pr\{C\}$, ... si $B$, $C$, ... son conjeturas, es decir, enunciados formulados claramente (quizá matemáticamente), de los cuales, no obstante, don Fulano de Tal no sabe en este momento si son ciertos o falsos. Sin embargo, si $A$ es cierto y don Fulano de Tal lo sabe, tendremos que $\Pr\{A\} = 1$. Si $A$ es falso y don Fulano de Tal lo sabe, tendremos que $\Pr\{A\} = 0$.

Me parece que nuestra ignoracia del valor numérico de $\Pr\{A\}$, $\Pr\{B\}$, ... no puede realmente perjudicarnos. De hecho, no nos conciernen aquí las opiniones personales de este señor. Lo que nos concierne son las reglas universales e impersonales de la inferencia plausible. Deseamos saber, en primer lugar, si existen tales reglas en absoluto, y luego, si el cálculo de probabilidad las revela (como Laplace y otros mantienen) o no. Por el momento, nosotros creemos que sí existen tales reglas, y que don Fulano de Tal, como persona sensata, razona de acuerdo con ellas, cualquiera grados de crédito $\Pr\{A\}$, $\Pr\{B\}$, $\Pr\{C\}$, ... pueda dar en este momento a los enunciados $A$, $B$, $C$, ... examinados. Por tanto, no puedo ver por qué nuestra ignorancia de los valores numéricos de $\Pr\{A\}$, $\Pr\{B\}$, $\Pr\{C\}$, ... nos perjudicaría.

Intentemos, en consecuencia, aplicar las reglas del cálculo de probabilidad a las credibilidades $\Pr\{A\}$, $\Pr\{B\}$, $\Pr\{C\}$, ... en cuanto interpretadas: fracciones positivas que miden grados de confianza de esa mítica o idealizada persona, don Fulano de Tal. Deseamos saber si, al hacerlo así, podemos obtener algo que sea razonablemente interpretado como una regla universal e impersonal de razonamiento plausible.

Nuestro intento puede fracasar, pero no vemos hasta el momento por qué fracasaría, y, en consecuencia, nos sentimos cautelosamente esperanzados *.

## 6. Examinando una consecuencia

Don Fulano de Tal está investigando una cierta conjetura $A$. Esta conjetura $A$ está claramente formulada, pero él no sabe si $A$ es cierta o no y desea saber en qué caso se halla: ¿$A$ es cierta o falsa? Observa una consecuencia $B$ de $A$ y considera satisfactorio el que

$$A \text{ implica } B.$$

Pero él no sabe si $B$ es verdadera o falsa, y algunas veces, cuando está cansado de investigar $A$, piensa dedicarse a la investigación de $B$.

Hemos considerado esta situación muchas veces, y ahora la reconsideramos a la luz del cálculo de probabilidad. Deseamos poner atención en tres credibilidades: $\Pr\{A\}$, $\Pr\{B\}$ y $\Pr\{A/B\}$. Don Fulano de Tal sabe muy bien que $A$ y $B$ no están probadas ni refutadas, pero él cree en ellas hasta cierto grado y este grado es expresado por $\Pr\{A\}$ y $\Pr\{B\}$, respectivamente. También $\Pr\{A/B\}$, el grado de creencia que podría tomar $A$ si $B$ fuese verdadera, juega un importante papel en sus deliberaciones.

No tenemos medio de asignar un valor numérico a cualquiera de estas credibilidades, aunque algunas veces podemos suponer en qué dirección un cambio en el conocimiento de don Fulano de Tal podría cambiar el valor de una o de otra. En cualquier caso, el cálculo de probabilidad da una relación entre ellas.

De hecho, por uno de los teoremas fundamentales sobre probabilidad (ver ejemplo 14.26 (2)), tenemos

$$\Pr\{A\}, \Pr\{B/A\} = \Pr\{B\} \Pr\{A/B\}.$$

* Que el cálculo de probabilidad trata primariamente con grados de creencia (confianza, confirmación, certidumbre...) y no con más o menos idealizadas frecuencias relativas es opinión de muchos autores, entre los que citaré dos: J. M. Keynes, *A treatise on probability* (cf., en especial, págs. 34, 66, 100), y B. de Finetti, «La prévision, ses lois logiques, ses sources subjectives», *Annales de l'Institut Henri Poincaré*, volumen 7 (1937), págs. 1-68. No hay espacio para explicar mis diferencias con estos autores, o las diferencias entre ellos, pero deseo expresar mi agradecimiento a ambos. El punto de vista adoptado aquí es semejante al de mi anterior publicación, aunque no igual: «Heuristic reasoning and the theory of probability», *American Math. Monthly*, volumen 48 (1941), págs. 450-465.

Pero ahora, puesto que $A$ implica $B$, $B$ debe ser verdadera si $A$ es verdadera, y así,

$$\Pr\{B/A\} = 1.$$

De aquí obtenemos

$$\text{(I)} \qquad \Pr\{A\} = \Pr\{B\}\,\Pr\{A/B\}.$$

Visualicemos los contenidos de esta ecuación.

(1) Don Fulano de Tal decidió investigar la consecuencia $B$ de su conjetura $A$. Sin embargo, no logró llegar a una conclusión en esta investigación, aunque a veces vio indicaciones de que $B$ podía ser cierta, y a veces indicaciones en contrario. Su confianza en $B$, que llamamos $\Pr\{B\}$, cae y se levanta, según los casos. Pero no vio nada que le hiciese cambiar sus puntos de vista sobre la relación entre $A$ y $B$ o sobre $\Pr\{A/B\}$. ¿Cómo influyó esto en $\Pr\{A\}$, su confianza en $A$?

La ecuación (I) muestra que $\Pr\{A\}$ *cambia en la misma dirección que* $\Pr\{B\}$, siempre que $\Pr\{A/B\}$ permanezcan invariables. Esto está de acuerdo con nuestras primeras observaciones, en especial con las de la sección 13.6. (Observemos que consideramos sólo la dirección del *cambio* averiguable a veces, y no la *cantidad*, imposible de saber con precisión.)

(2) Don Fulano de Tal logró probar $B$, que es una consecuencia de la conjetura $A$ que originalmente investigó. Antes de probar $B$ tenía alguna confianza en $B$, cuyo grado hemos representado por $\Pr\{B\}$; tenía también alguna confianza en $A$, cuyo grado llamamos $\Pr\{A\}$. Algunas veces consideró $\Pr\{A/B\}$, la confianza que tomaría $A$ tras una prueba de $B$. Tras la prueba de $B$, su confianza en $B$ alcanzó el valor máximo 1, y su confianza en $A$ fue, por supuesto, $\Pr\{A/B\}$. (Sustituyendo $\Pr\{B\}$ por 1 en la ecuación (I), nos vemos conducidos al nuevo valor de la credibilidad de $A$.) Suponemos aquí que su punto de vista sobre la relación entre $A$ y $B$ y su evaluación de $\Pr\{A/B\}$ permanece invariable.

Observando que $0 < \Pr\{B\} < 1$, derivamos de la ecuación (I) la desigualdad

$$\text{(II)} \qquad \Pr\{A\} < \Pr\{A/B\}.$$

Ahora, $\Pr\{A\}$ y $\Pr\{A/B\}$ representan la credibilidad de $A$ antes y después de la prueba de $B$, respectivamente. Por tanto, la desigualdad (II) es la expresión formal de un principio con el que nos encontramos a menudo: ***la verificación de una consecuencia hace a la conjetura más digna de crédito***; cf. la sección 12.1, por ejemplo.

(3) Pero aún podemos aprender más de la ecuación (I), que escribimos en la forma:

$$\text{(III)} \qquad \Pr\{A/B\} = \frac{\Pr\{A\}}{Pr\{B\}}.$$

El lado izquierdo, la credibilidad de $A$ tras la verificación de $B$, se expresa en términos de la confianza que el investigador tenía en $A$ y $B$, respectivamente, antes de tal verificación. Comparemos varios casos de verificación con éxito de una consecuencia. Estos casos tienen una circunstancia en común: la misma confianza $\Pr\{A\}$ se dio a la conjetura $A$ (a que la investigación apunta) antes de la verificación de su consecuencia $B$. Pero los casos difieren en otro aspecto: la consecuencia $B$ (eventualmente verificada) fue esperada con más confianza en algunos casos y con menos confianza en otros. Es decir, consideramos $\Pr\{A\}$ como constante y $\Pr\{B\}$ como variable. ¿Cómo la variación de $\Pr\{B\}$ influye en el peso de la evidencia que resulta de la verificación de la consecuencia $B$?

Prestemos la atención debida a los casos extremos. Puesto que $B$ es una consecuencia de $A$, $B$ es verdadera cuando $A$ es verdadera, y así, $\Pr\{B\}$, la credibilidad $B$, no puede ser menor que $\Pr\{A\}$, la credibilidad de $A$. Por otro lado, ninguna credibilidad puede exceder a la certeza: $\Pr\{B\}$ no puede ser mayor de 1. Hemos determinado los límites en que se contiene $\Pr\{B\}$.

$$\Pr\{A\} \leqq \Pr\{B\} < 1.$$

El límite más bajo es alcanzado, cuando no sólo $A$ implica a $B$, sino que también $B$ implica a $A$, de manera que las dos aserciones $A$ y $B$ son equivalentes, se mantienen y caen juntas, y son, claro está, igualmente dignas de crédito. El límite más alto 1 no puede ser realmente alcanzado: si fuera alcanzado, $B$ sería cierta antes de ser investigada y no hemos incluido este caso en nuestra consideración. No obstante, podemos aproximarnos al límite más alto: $B$ puede ser casi cierta antes de ser examinada. ¿Cómo cambia la evidencia resultante de la verificación de $B$ cuando $\Pr\{B\}$ varía entre sus límites extremos?

La evidencia es más fuerte cuando $\Pr\{A/B\}$, la nueva confianza en $A$ que resulta de la verificación de la consecuencia $B$, es más grande. Esto se ve desde la relación (III) que

cuando $\Pr\{B\}$ disminuye de 1 a $\Pr\{A\}$

$\Pr\{A/B\}$ aumenta de $\Pr\{A\}$ a 1.

Este enunciado expresa en un nuevo lenguaje un punto que hemos reconocido antes (sección 12.3): *el aumento de nuestra confianza en una conjetura debida a la verificación de una de sus consecuencias varía inversamente a la credibilidad de la consecuencia antes de ta verificación.* Cuanto más inesperada es una consecuencia, mayor peso da su verificación. La verificación de las consecuencias más sorprendentes son las más convincentes, mientras que la verificación de una consecuencia que no nos ofrece muchas dudas ($\Pr\{B\}$ casi 1) tiene poco valor como evidencia.

(4) La situación examinada puede ser vista desde otro punto de vista. Utilizando algunas simples reglas del cálculo de probabilidad (cf. ejemplo 14.26, fórmulas (4), (2), (3), en este orden) obtenemos

$$\begin{aligned}\Pr\{B\} &= \Pr\{AB\} + \Pr\{\bar{A}B\} \\ &= \Pr\{A\}\Pr\{B/A\} + \Pr\{\bar{A}\}\Pr\{B/\bar{A}\} \\ &= \Pr\{A\} + [1 - \Pr\{A\}]\Pr\{B/\bar{A}\}.\end{aligned}$$

Al pasar a la última línea, también utilizamos $\Pr\{B/A\} = 1$, que expresa que $B$ es una consecuencia de $A$. Al sustituir por $\Pr\{B\}$ el valor derivado, obtenemos de (III)

$$\text{(IV)} \qquad \Pr\{A/B\} = \frac{\Pr 0A 0}{\Pr\{A\} + [1 - \Pr\{A\}]\Pr\{B/\bar{A}\}}.$$

Asumimos como antes que $\Pr\{A\}$ es constante; es decir, examinamos varios casos en que la confianza en $A$, la conjetura examinada, fue la misma antes de comprobar la consecuencia $B$ de $A$. Pero $\Pr\{A/B\}$, la credibilidad de $A$ tras la verificación de $B$, aún depende de $\Pr\{B/\bar{A}\}$, la credibilidad de $B$ (antes de la verificación, claro es) vista bajo la asunción de que $A$ no es cierta. Y $\Pr\{B/\bar{A}\}$ puede variar; en efecto, puede tomar cualquier valor entre 0 y 1. Ahora, según nuestra fórmula (III),

cuando $\Pr\{B/\bar{A}\}$ disminuye de 1 a 0

$\Pr\{A/B\}$ aumenta de $\Pr\{A\}$ a 1.

Este enunciado expresa en un nuevo lenguaje un punto que hemos examinado antes en la sección 13.10. Veamos los casos extremos. Si $B$ sin $A$ es difícil de creer ($\Pr\{B/\bar{A}\}$ casi 0), la verificación de la consecuencia $B$ trae la conjetura $A$ cerca de la certeza. Por otro lado, la verificación de una consecuencia $B$ que nosotros apenas dudamos aun cuando $A$ fuese falsa ($\Pr\{B/\bar{A}\}$ casi 1) añade poco a nuestra confianza en $A$.

## 7. Examinando un fundamento posible

Tras el amplio y precavido examen de la sección anterior podemos proceder a una consideración un poco más rápida de situaciones semejantes.

Aquí hay una situación tal: el objetivo de nuestra investigación es una cierta conjetura $A$. Observamos un fundamento posible para $A$, es decir, una proposición $B$ de la que $A$ se sigue:

$A$ está implicada en $B$.

Empezamos la investigación de $B$. Si logramos probar $B$, $A$ también quedaría probada. Sin embargo, $B$ resulta ser falsa. ¿Cómo afecta la falsedad de $B$ a nuestra confianza en $A$?

El cálculo de probabilidad contesta a esta pregunta. Puesto que $A$ está implicada en $B$,

$$\Pr\{A/B\} = 1.$$

Combinemos esto con algunas fórmulas básicas (ver ejemplo 14.26 (4), (2), (3)):

$$\begin{aligned}\Pr\{A\} &= \Pr\{AB\} + \Pr\{A\bar{B}\} \\ &= \Pr\{B\}\Pr\{A/B\} + \Pr\{\bar{B}\}\Pr\{A/\bar{B}\} \\ &= \Pr\{B\} + (1 - \Pr\{B\})\Pr\{A/\bar{B}\}.\end{aligned}$$

Obtenemos de aquí que

$$\text{(I)} \qquad \Pr\{A/\bar{B}\} = \frac{\Pr\{A\} - \Pr\{B\}}{1 - \Pr\{B\}}.$$

El lado izquierdo que representa la credibilidad de $A$ tras $B$ (que es un fundamento posible de $A$) ha sido refutado. El lado derecho se refiere a la situación antes de la refutación de $B$. Por cierto que este lado derecho puede transformarse de modo que la ecuación (I) aparezca en la forma

$$\Pr\{A/\bar{B}\} = \Pr\{A\} - \Pr\{B\}\frac{1 - \Pr\{A\}}{1 - \Pr\{B\}}$$

y de aquí vemos que

$$\text{(II)} \qquad \Pr\{A/\bar{B}\} < \Pr\{A\}.$$

Ambos lados de esta desigualdad representan la credibilidad de la conjetura $A$, el lado izquierdo después de la refutación de $B$, el lado derecho antes de dicha refutación. Por tanto, la desigualdad (III) expresa una regla: *nuestra confianza en una conjetura sólo puede disminuir cuando un fundamento posible para la conjetura ha sido refutado.* (Cf. sección 13.2.)

Sin embargo, aún podemos aprender más de la ecuación (I). Consideremos $\Pr\{A\}$ como constante y $\Pr\{B\}$ como variable. Esto es, examinemos varios casos que defieren en un importante aspecto: en el grado de nuestra confianza en $B$, que puede ser muy pequeña, pero no puede ser arbitrariamente mayor: nunca puede exceder a nuestra confianza en $A$, puesto que si $B$ es verdadera $A$ es también verdadera. (No obstante, $A$ podría ser aún verdadera, incluso siendo $B$ falsa.) Y así, determinamos los valores extremos entre los que puede variar $\Pr\{B\}$:

$$0 < \Pr\{B\} \leqq \Pr\{A\}.$$

Por la ecuación (I) vemos que

$$\text{a medida que } \Pr\{B\} \text{ aumenta desde 0 a } \Pr\{A\}$$

$$\Pr\{A/\bar{B}\} \text{ disminuye desde } \Pr\{A\} \text{ a 0.}$$

Es decir, *cuanto más confianza pongamos en una posible base para nuestra conjetura, mayor será la pérdida de fe en ésta cuando dicha base sea refutada.*

## 8. Examen de conjeturas en conflicto

Consideramos ahora otra situación: tenemos dos conjeturas, $A$ y $B$, en conflicto. Cuando decimos que $A$ está en conflicto con $B$ o que

$$A \text{ es incompatible con } B$$

entendemos por ello que la verdad de una de ellas implica la falsedad de la otra. De hecho, estamos, en primer lugar, ocupándonos de $A$ y hemos empezado a investigar $B$ porque pensamos que la investigación de ésta podría arrojar alguna luz sobre $A$. Una prueba de $B$ refutaría a $A$. Sin embargo, logramos refutar a $B$. ¿Cómo afecta este resultado a nuestra confianza en $A$?

Dejemos al cálculo de probabilidades dar la respuesta. Empecemos por expresar en el lenguaje de este cálculo que $A$ y $B$ son incompatibles,

Esto significa, en otras palabras, que $A$ y $B$ no pueden ser ambas verdaderas, y, por tanto,

$$\Pr\{AB\} = 0.$$

Ahora concluimos, según nuestras fórmulas básicas (cf. ejemplo 14.26 (4), (2), (3)):

$$\begin{aligned}\Pr\{A\} &= \Pr\{AB\} + \Pr\{A\bar{B}\} \\ &= \Pr\{A\bar{B}\} \\ &= \Pr\{\bar{B}\}\Pr\{A/\bar{B}\} \\ &= (1 - \Pr\{B\})\Pr\{A/\bar{B}\}.\end{aligned}$$

Lo que arroja, finalmente,

$$\text{(I)} \qquad \Pr\{A/\bar{B}\} = \frac{\Pr\{A\}}{1 - \Pr\{B\}}.$$

La ecuación (I) implica, obviamente, la desigualdad:

$$\text{(II)} \qquad (\Pr\{A/\bar{B}\} < \Pr\{A\}).$$

El lado izquierdo de la ecuación se refiere a la situación después de haber sido refutada $B$; el lado derecho, a la situación antes de esa refutación. Por tanto, podemos leer (II) como sigue: *nuestra confianza en una conjetura puede sólo aumentar cuando una conjetura rival incompatible ha sido refutada*]. (Cf. sección 13.3.)

Pero aún podemos aprender más de la ecuación (I). Consideremos $\Pr\{A\}$ como constante y $\Pr\{B\}$ como variable. Determinemos los límites entre los que puede variar $\Pr\{B\}$. Naturalmente, $\Pr\{B\}$ puede ser tan pequeña como se quiera, pero no arbitrariamente grande; en efecto, no puede nunca exceder a $\Pr\{\bar{A}\}$. Si $B$ es correcta, $\bar{A}$ es correcta *a fortiori*. Puesto que $\Pr\{\bar{A}\}$ es igual a $1 - \Pr\{A\}$,

$$0 < \Pr\{B\} \leqq 1 - \Pr\{A\}.$$

Por la ecuación (I) vemos que

a medida que $\Pr\{B\}$ aumenta desde 0 a $1 - \Pr\{A\}$
$\Pr\{A/\bar{B}\}$ aumenta desde $\Pr\{A\}$ a 1.

Es decir, *cuanta más confianza ponemos en una rival incompatible de nuestra conjetura, mayor será la fe en nuestra conjetura cuando su rival haya sido refutada.*

## 9. Examinando varias consecuencias en sucesion

Consideremos ahora la siguiente importante situación: el objeto de nuestro trabajo es cierta conjetura $A$. Por el momento, no vemos cómo vamos a decidir si es verdadera o no. Sin embargo, vemos varias consecuencias $B_1$, $B_2$, $B_3$, ... de $A$:

$A$ implica a $B_1$, $A$ implica a $B_2$, $A$ implica a $B_3$, ... Las consecuencias $B_1$, $B_2$, $B_3$, son más accesibles que $A$ misma y nos proponemos examinarlas una después de otra. (Este es el procedimiento típico de las ciencias naturales: no tenemos medios de examinar una ley general $A$ en sí misma, y, por tanto, la examinamos comprobando varias consecuencias $B_1$, $B_2$, $B_3$, ...) Hemos ya examinado las consecuencias $B_1$, $B_2$, ... $B_n$, y al verificarlas a todas: $B_1$, $B_2$, ..., $B_n$, las hemos hallado correctas. Ahora comprobamos la próxima consecuencia $B_{n+1}$: ¿Cómo efectará nuestra confianza en $A$?

Al objeto de ver la contestación a la luz del cálculo de probabilidad empecemos a partir de una regla general de este cálculo (ver ejemplo 14.26 (5)):

$$\Pr\{A/H\}\Pr\{B/HA\} = \Pr\{B/H\}\Pr\{A/HB\}.$$

Hagamos $B = B_{n+1}$. Ahora, puesto que $B_{n+1}$ es una consecuencia de $A$,

$$\Pr\{B/HA\} = \Pr\{B_{n+1}/HA\} = 1,$$

y así, encontramos que

$$\Pr\{A/H\} = \Pr\{B_{n+1}/H\}\Pr\{A/HB_{n+1}\}.$$

Hagamos $H = B_1, B_2 \ldots B_n$ y obtendremos está fórmula decisiva:

$$\text{(I)} \qquad \Pr\{A/B_1 \ldots B_n\} = \Pr\{B_{n+1}/B_1 \ldots B_n\}\Pr\{A/B_1 \ldots B_n B_{n+1}\}.$$

Al objeto de comprender (I) correctamente, tenemos que comprobar que

$$\Pr\{A/B_1 \ldots B_n\} \text{ y } \Pr\{B_{n+1}/B_1 \ldots B_n\}$$

denota las credibilidades $A$ y $B_{n+1}$, respectivamente, *después* de haber sido verificadas $B_1$, $B_2$, ... $B_n$, pero, naturalmente, *antes de* que $B_{n+1}$ haya sido verificada;

$\Pr\{A/B_1 \ldots B_n B_{n+1}\}$ denota la credibilidad de $A$ después de la verificación de sus $n+1$ consecuencias, $B_1$, $B_2$, ... $B_n$ y $B_{n+1}$.

Mantengamos estos significados en la cabeza, y después podemos

leer (I) como una proposición precisa y fecunda sobre razonamiento inductivo.

Primero pongamos nuestra atención en $\Pr\{B_{n+1}/B_1 \ldots B_n\}$; el valor de esta credibilidad será, en la mayoría de los casos, menos de 1, e igual a 1 sólo si la corrección de $B_1$, $B_2$, ... $B_n$ hace la corrección de $B_{n+1}$ cierta, es decir, si $B_1$, $B_2$, ... y $B_n$ implican conjuntamente $B_{n+1}$. Si éste *no* es el caso, podemos derivar de (I) la desigualdad:

$$\text{(II)} \qquad \Pr\{A/B_1 \ldots B_n\} < \Pr\{A/B_1 \ldots B_n B_{n+1}\}.$$

Esto es, *la verificación de una nueva consecuencia aumenta nuestra confianza en la conjetura, a menos que la nueva consecuencia esté implícita en consecuencias formalmente verificadas.*

Escribamos la ecuación (I) de la forma:

$$\text{(III)} \qquad \Pr\{A/B_1 \ldots B_n B_{n+1}\} = \frac{\Pr\{A/B_1 \ldots B_n\}}{\Pr\{B_{n+1}/B_1 \ldots B_n\}}.$$

El lado izquierdo se refiere a la situación después de la confirmación de $B_{n+1}$; el derecho a la situación de antes de dicha confirmación. Vamos a considerar la relación de $A$ con $B_1$, $B_2$, ... $B_n$ como fija, y la relación de $B_{n+1}$ con $B_1$ $B_2$ ... $B_n$ como variable. Entonces, podemos leer (III) como sigue: *el aumento de nuestra confianza producido por la confirmación de una nueva consecuencia*(o el peso de la evidencia proporcionado por esta confirmación) *varía en proporción inversa a la credibilidad de la nueva consecuencia, apreciada* (antes de su confirmación, claro está) *a la luz de las consecuencias previamente verificadas.*

Podemos expresar esta misma regla en otras palabras. Cuando empezamos a probar la consecuencia $B_{n+1}$ de nuestra conjetura $A$ nos enfrentamos con la posibilidad de que $B_{n+1}$ resulte falsa, en cuyo caso $A$ será refutada. En vista de las consecuencias primeramente verificadas $B_1$, $B_2$, ... $B_n$ la posibilidad de refutar $A$, no habiendo podido probar $B_{n+1}$, es fuerte cuando $\Pr\{B_{n+1}/B_1 \ldots B_n\}$ es pequeña. Por tanto, podemos leer (III) del modo siguiente: *la consecuencia que, juzgada a la luz de las verificaciones precedentes, significa la mejor posibilidad de refutar la conjetura propuesta revelará la evidencia inductiva más fuerte, si es confirmada a despecho de los temores en contra.* O, más brevemente: «a más peligro más honor». Si una conjetura escapa al peligro de ser refutada será estimada en proporción al riesgo implicado.

Desde el principio de nuestra discusión hemos considerado la evidencia inductiva suministrada por la verificación sucesiva de varias consecuencias de una conjetura propuesta. Los casos extremos fueron

los más notables. Examinémolos una vez más (añadiendo un poco de color) y dirijamos la atención al momento en que, una vez verificadas las consecuencias $B_1$, $B_2$, ... $B_n$ de una conjetura $A$ empezamos a escudriñar una nueva consecuencia $B_{n+1}$.

La nueva consecuencia en examen, $B_{n+1}$, puede parecer «poco diferente» de las consecuencias verificadas anteriormente $B_1$, $B_2$, ... $B_n$. Un caso semejante no es demasiado interesante. Confiadamente esperamos (por analogía, sin duda) que $B_{n+1}$ se verificará como las otras consecuencias (es decir, $\Pr\{B_{n+1}/B_1 \ldots B_n\}$ está próxima a 1, que es su máximo). Apenas esperamos que la investigación de $B_{n+1}$ revele algún aspecto nuevo o que desaparezca la conjetura $A$, pero también cuando $B_{n+1}$ sea, por fin, verificada, la evidencia ganada por $A$ no habrá sido mucha.

Por otro lado, la nueva consecuencia, que estamos examinando, $B_{n+1}$, puede parecer «muy diferente» de las consecuencias $B_1$, $B_2$, ... $B_n$ verificadas en primer lugar. Un caso semejante puede ser interesante. La analogía con $B_1$, $B_2$, ... $B_n$ nos da poca razón para esperar que $B_{n+1}$ sea verificada ($\Pr\{B_{n+1}/B_1 \ldots B_n\}$ está próxima a su mínimo). Nos damos cuenta de que la investigación de $B_{n+1}$ puede trastornar la conjetura $A$, aunque también es una oportunidad para revelar algún aspecto nuevo, y cuando $B_{n+1}$ sea, por fin, verificada, la ganancia en evidencia de $A$ puede ser considerable.

El lector deberá revisar algunos de nuestros primeros ejemplos y discusiones. (Cf. secciones 3.1-3.7, capítulo VI, secciones 10.1, 12.2 y 13.11, y varios otros pasajes.) Tras la debida comparación, la ecuación (III) de la sección presente puede parecer la expresión más concisa y precisa del principio que implica. De cualquier modo, si usted puede ver la orientación de la ecuación (III) en alguno de nuestros ejemplos ha dado un buen paso para aclarar sus ideas sobre un importante tema.

## 10. Sobre la evidencia circunstancial

Consideremos ahora una situación, que hemos encontrado ya al tratar con el razonamiento en materias judiciales: estamos examinando una conjetura $A$. (Esta conjetura puede ser una acusación presentada por el fiscal.) Nosotros (el jurado) tenemos que hallar si $A$ es verdadera o no. El fiscal somete una circunstancia $B$, la cual está relacionada con la conjetura $A$, de modo que

$B$ con $A$ es más digna de crédito que sin $A$.

En el curso del proceso esta circunstancia $B$ resulta confirmada tan fuertemente que podemos considerarla como un hecho probado. (Quizá $B$ no ha sido recusada por la defensa.) ¿Cómo afecta todo esto a nuestra creencia en $A$?

El cálculo de probabilidades contestará a esta pregunta. El presupuesto esencial que concierne a la conexión entre $A$ y $B$ está expresado por la desigualdad

$$\text{(I)} \qquad \Pr\{B/A\} < \Pr\{B/\bar{A}\};.$$

Por esta fórmula de la teoría de la probabilidad (cf. ejemplo 14.26,)

$$\Pr\{A\}\Pr\{B/A\} = \Pr\{B\}\Pr\{A/B\},$$
$$\Pr\{B\} = \Pr\{A\}\Pr\{B/A\} + (1 - \Pr\{A\})\Pr\{B/\bar{A}\}.$$

Combinando estas dos obtenemos

$$\text{(II)} \qquad \frac{\Pr\{A\}}{\Pr\{A/B\}} = \Pr\{A\} + (1 - \Pr\{A\})\frac{\Pr\{B/\bar{A}\}}{\Pr\{B/A\}}.$$

Utilizando (I), concluimos de (II) que

$$\text{(III)} \qquad \Pr\{A\} < \Pr\{A/B\}.$$

Ambos lados de esta desigualdad representan la credibilidad de la conjetura $A$, el lado izquierdo antes de la verificación de la circunstancia $B$, el lado derecho después de dicha verificación. Por tanto, la desigualdad (III) expresa una regla: *si cierta circunstancia es más digna de crédito con una determinada conjetura que sin ella, la prueba de esa circunstancia sólo puede aumentar la credibilidad de dicha conjetura.* (Cf. sección 13.13.)

Pero aún podemos aprender más de la ecuación (II). Consideremos $\Pr\{A\}$ y $\Pr\{B/A\}$ como constantes, pero $\Pr\{B/\bar{A}\}$ como variable. Luego $\Pr\{A/B\}$ depende de $\Pr\{B/\bar{A}\}$:

cuando $\Pr\{B/\bar{A}\}$ disminuye de $\Pr\{B/A\}$ a 0
$\Pr\{A/B\}$ aumenta de $\Pr\{A\}$ a 1.

Es decir, *cuanto menos digna de crédito parece una circunstancia sin una conjetura determinada más aumentará la prueba de esa circunstancia la probabilidad de la conjetura dicha.* En la sección 13.13, conducidos por la consideración de ejemplos, llegamos muy cerca de esta regla.

Una fuerte evidencia judicial resulta de varias coincidencias que apuntan todos a la misma conclusión; ver sección 13.13 (4). Si hay

varias circunstancias $B_1$, $B_2$, $B_3$, ..., cada una de las cuales es más digna de crédito con $A$ que sin $A$, y todas son probadas sucesivamente, la evidencia de $A$ aumenta en cada paso. La cantidad de evidencia adicional que resulta de una circunstancia nuevamente probada depende de varios puntos. Una nueva circunstancia que es muy diferente de las circunstancias previamente examinadas (un nuevo testimonio que es completamente distinto de los testimonios presentados antes) tiene peso especial. Para expresar estos puntos podríamos desarrollar fórmulas relacionadas con las presentadas en esta sección como las fórmulas desarrolladas en la sección 9 están relacionadas con las presentadas en la sección 6.

## EJEMPLOS Y COMENTARIOS

**1.** Examinar la situación discutida en la sección 13.8 por el cálculo de probabilidad.

**2.** Examinar el patrón encontrado en la solución del ejemplo 13.8 por el cálculo de probabilidad. ¿Puede usted justificarlo?

**3.** Reexaminar el ejemplo 13.10 por el cálculo de probabilidad.

**4.** *Probabilidad y credibilidad.* La peculiar aplicación «no-cuantitativa» del cálculo de probabilidad, examinada en la sección 5-10, fue delineada para elucidar ciertos patrones de razonamiento plausible. Estos patrones son principalmente sugeridos por razonamientos heurísticos sobre conjeturas matemáticas. ¿Podemos aplicar el cálculo de probabilidad a ejemplos de otra clase del mismo modo?

(1) $A_n$ denota la conjetura de que el dado «justo» que estoy volteando sacará $n$ puntos ($n = 1, 2, \ldots 6$). Estas conjeturas $A_1, A_2, \ldots A_6$ son de una clase que no deseábamos examinar especialmente en la sección 5-10. Sin embargo, intentemos ahora tratarlas de la misma manera: consideremos sus crecibilidades $\mathrm{Pr}\{A_1\}$, $\mathrm{Pr}\{A_2\}$, ... $\mathrm{Pr}\{A_6\}$ y apliquémosle el cálculo de probabilidad. Como $A_1, A_2, \ldots A_6$ son mutuamente exclusivas y agotan todas las posibilidades,

$$\mathrm{Pr}\{A_1\} + \mathrm{Pr}\{A_2\} + \ldots + \mathrm{Pr}\{A_6\} = 1.$$

Puesto que el dado que utilizamos es «justo», sus caras son intercambiables, ninguna de ellas es preferible a otra, y por ello nos sentimos impulsados a asumir que

$$\mathrm{Pr}\{A_1\} = \mathrm{Pr}\{A_2\} = \ldots = \mathrm{Pr}\{A_6\}.$$

Combinando ésta con la ecuación anterior, obtenemos que

$$\Pr\{A_1\} = \Pr\{A_2\} = \ldots = \Pr\{A_6\} = 1/6,$$

y así, nos vemos conducidos a atribuir un valor numérico definido a las credibilidades consideradas. Don Fulano de Tal (nuestro amigo de la sección 5) yo creo que se vería conducido a la misma conclusión.

La credibilidad de la conjetura $A_1$ resultó tener el mismo valor numérico que la probabilidad de que un dado «justo» saque un punto. Pero esto no nos sorprende en absoluto: admitimos las mismas reglas y asumimos la misma intercambiabilidad (o simetría) al calcular credibilidades y probabilidades. (El lector no olvidará, naturalmente, que credibilidad y probabilidad son definidas de modo bastante diferente.)

(2) El anterior argumento se aplica a muchos otros casos.

La superficie de un círculo es dividida en seis sectores iguales por tres líneas rectas que pasan por su centro. Una gota de lluvia está para caer en la superficie circular. Llamamos $A_1$ la afirmación de que la gota de lluvia caerá dentro del primer sector, $A_2$ la afirmación de que caerá en el segundo, y así sucesivamente. Esta situación es esencialmente la misma que la examinada en (1) y el resultado es, naturalmente, también el mismo: cada una de las seis afirmaciones $A_1$, $A_2$, ... $A_6$ tienen la misma credibilidad 1/6.

La generalización es obvia: no necesitamos adherirnos al número 6 o a las gotas de lluvia. Podemos dividir el círculo en $n$ sectores y considerar cualquier otra clase de suceso azaroso. Podemos pasar también del área del círculo a su periferia y, así, llegamos a la siguiente situación: un agente azaroso que no tiene ninguna preferencia particular por cualquiera de los puntos de la periferia del círculo elegirá, sin embargo, un punto de esta periferia. Alguien conjetura que dicho punto estará en cierto arco. La credibilidad de esta conjetura es la razón de la longitud del arco respecto a la periferia total.

Del círculo podemos pasar a la esfera. Asumamos, en pro de la simplicidad, que la Tierra es una esfera perfecta y que los meteoritos que chocan contra ella no tienen preferencia por ninguna dirección particular. Alguien conjetura que el próximo meteorito caerá en cierta región. ¿Cuál es la credibilidad de esta conjetura? El procedimiento matemático puede ser más complicado si lo realizamos con todo detalle, pero el resultado es exactamente tan intuitivo como el del círculo: la credibilidad deseada es la razón del área de esa región respecto al área de toda la superficie esférica.

No necesitamos entrar en nuevas situaciones análogas o más generales. Sin embargo, veamos dos casos particulares.

La credibilidad de la conjetura de que un punto elegido al azar en la periferia de un círculo estará dentro de la distancia de un grado, a partir de un punto dado de dicha periferia, es 1/180.

La credibilidad de la conjetura de que el próximo meteorito que caiga sobre la Tierra lo hará dentro de la distancia de un grado a partir del centro de Nueva York es 0,00007615. (Esta es la razón numérica de un pequeño casquete esférico respecto de toda la superficie esférica.)

El círculo y la esfera poseen una gran simetría: una rotación conveniente que no cambie la posición de la figura como un todo puede transportar un punto de la figura desde cualquier posición dada sobre ésta a cualquiera otra posición dada. Esta simetría no es bastante en sí misma para justificar los resultados precedentes. Al derivarlos tenemos que *asumir* la simetría «física», la intercambiabilidad de dos puntos cualesquiera de la figura geométrica en relación con el agente físico que estamos considerando.

Las credibilidades calculadas no pueden pretender mucha novedad: coinciden con probabilidades correspondientes conocidas durante largo tiempo y aplicadas con éxito a una variedad de fenómenos de masas.

Las credibilidades calculadas pueden parecer incapaces de interferir con las credibilidades de que nos hemos ocupado primeramente, pero no estamos demasiado seguros de esto.

(3) Kepler conoció sólo seis planetas que giraban alrededor del Sol e incluso ideó un argumento geométrico que justificaba el número exacto de seis; ver sección 11.5. No obstante, el telescopio no escuchó su argumento. En 1781, unos ciento cincuenta años después de la muerte de Kepler, el astrónomo Herschel observó una estrella que se movía lentamente y supuso que era un cometa, pero resultó ser el séptimo planeta, Urano, que giraba más allá de la órbita de Saturno. Entre los años 1801-1806 se descubrieron de modo semejante cuatro pequeños planetas (Ceres, Alas, Palas, Juno y Vesta) que giraban entre las órbitas de Marte y Júpiter. Más tarde se han descubierto cientos de estos planetas menores.

Sobre la base de la teoría de Newton los astrónomos intentaron calcular los movimientos de estos planetas. No tuvieron mucho éxito con el planeta Urano; las diferencias entre teoría y observación parecían exceder los límites admisibles de error. Algunos astrónomos sospecharon que estas desviaciones podían deberse a la atracción de un planeta que se movía más allá de la órbita de Urano, y el astrónomo francés Leverrier investigó esta conjetura más profundamente que sus colegas. Examinando las distintas explicaciones propuestas encontró que había sólo una que podía contar para las irregularidades observadas en el movimiento de Urano: la existencia de un planeta ultrauraniano. Intentó calcular

la órbita de ese hipotético planeta a partir de las irregularidades de Urano. Finalmente, Leverrier consiguió asignar una posición determinada en el cielo al hipotético planeta. Escribió sobre esto a otro astrónomo cuyo observatorio era el mejor equipado para examinar aquella porción del cielo. La carta llegó el 23 de septiembre de 1846, y en la tarde del mismo día se encontró un nuevo planeta a un grado del punto señalado por Leverrier. Era un gran planeta ultrauraniano que tenía, aproximadamente, la masa y órbita predichas por Leverrier.

(4) La teoría que hace posible dar una predicción tan extraordinaria debe ser una teoría maravillosa. Esta puede ser nuestra primera impresión. Vamos a tratar de aclarar esta impresión por el cálculo de probabilidades.

$T$ denota la teoría que sustenta los cálculos astronómicos: es la teoría de Newton, que consta de sus leyes de mecánica y de su ley de la gravitación universal.

$N$ denota la aserción de Leverrier: en la fecha dada habrá un nuevo planeta con tal y tal masa y tal y tal órbita (aproximadamente), en las cercanías de tal y tal punto del cielo. Más precisamente, $N$ denota esta parte de la aserción de Leverrier que ha sido *verificada* por observaciones subsecuentes.

$\Pr\{T\}$ denota el grado de confianza colocado en la teoría $T$ a la vista de los hechos conocidos por los astrónomos antes de la fecha del descubrimiento del nuevo planeta.

$\Pr\{T/N\}$ denota el grado de confianza debido a la misma teoría $T$ cuando la verificación de la predicción de Leverrier $N$ se ha añadido a los hechos antes conocidos.

Creemos que la verificación de la predicción de Leverrier aumenta el crédito de la teoría $T$, y, así, sospechamos que $\Pr\{T/N\}$ es mayor que $\Pr\{T\}$. En efecto, cncontramos (por el ejemplo 14.26 (2)) que

$$\frac{\Pr\{T/N\}}{\Pr\{T\}} = \frac{\Pr\{N/T\}}{\Pr\{N\}}.$$

$\Pr\{N\}$ es la credibilidad de la aserción de Leverrier $N$ «en sí misma», es decir, sin referencia a la verdad o falsedad de la teoría $T$; el estado de conocimiento subyacente es el mismo que asumimos en la evaluación de $\Pr\{T\}$. Esta $\Pr\{N\}$ debe ser extremadamente pequeña: Si ignoráramos la teoría de $T$ de Newton ¿qué motivo podría habernos dado la sospecha de que existe un planeta ultrauraniano de tales y tales precisas propiedades cerca de un determinado punto del cielo?

$\Pr\{N/T\}$ es la credibilidad de la aserción $N$ de Leverrier a la luz de

su cálculo basado en la teoría $T$, y por eso es un concepto muy diferente de $\Pr\{N\}$. Quizá Leverrier no mostró bastante conclusivamente que la existencia de un planeta ultrauraniano con tales y tales propiedades es la única aplicación compatible con la teoría $T$. Sin embargo, estuvo bastante cerca de mostrar esto, y, así, $\Pr\{N/T\}$ no puede estar demasiado lejos de la certeza.

Para resumir podemos considerar $\Pr\{N/T\}$ como una fracción proporcionada, no muy lejos de 1, y $\Pr\{N\}$ como una fracción muy pequeña que se aproxima a 0. No obstante, si pensamos así, la razón $\Pr\{N/T\}/\Pr\{N\}$ parece muy grande, y así debe parecer la razón $\Pr\{T/N\}\Pr\{T\}$, que tiene el mismo valor que la anterior, según la ecuación expresada más arriba. Por tanto, $\Pr\{T/N\}$, nuestra confianza en la teoría $T$ después de la verificación de la predicción de Leverrier aparece mucho mayor que $\Pr\{T\}$, nuestra confianza en la misma teoría antes de dicha verificación.

(4) La consideración anterior permanece dentro de los límites a los que cautelosamente nos hemos restringido en la sección 5. Pero vamos a dejar de lado la precaución por un momento y a permitirnos una aventurada estimación, hecha a *grosso modo*.

La predicción $N$ de Leverrier contiene muchos detalles, de los que tomamos sólo uno: el nuevo planeta estaría dentro de tal y tal punto del cielo. En efecto, se encontró a un lado del punto indicado (a una distancia de 52′). No obstante, la probabilidad de que un punto elegido al azar en la esfera esté a un grado del lugar señalado puede ser calculada por sencillos supuestos, como establecimos arriba, en (2). Encontramos que $\Pr\{N\}$ es mucho más sencilla que 0,00007615. Escasamente podemos considerar $\Pr\{N/T\}$ exactamente igual a 1. Pero podemos suponer que haciendo

$$\Pr\{N\} = 0.00007615, \qquad \Pr\{N/T\} = 1,$$

sobreestimamos la credibilidad del primero mucho más que la del segundo. Y así, llegamos a la desigualdad

$$\frac{\Pr\{T/N\}}{\Pr\{T\}} < \frac{1}{0,00007615} = 13131.$$

Naturalmente, tal estimación es objetable. Puede haber razones de analogía completamente independientes de la teoría $T$ de Newton que sugieran que un nuevo planeta tiene más probabilidades de estar cerca del plano de la órbita de la Tierra que lejos de él. Si pensamos así sus-

tituiremos 0.00007615 por una fracción mayor, pero que sea menor que

$$1/180 = 0{,}005556;$$

cf. (2).

Se pueden argüir tales estimados sin fin. Por ejemplo, puesto que $\Pr\{T/N\}$ es ciertamente menor que 1, la desigualdad propuesta implica que

$$\Pr\{T\} < 0{,}00007615.$$

Nos podemos sentir tentados de considerar esto como una refutación, como una *reductio ad absurdum* de la desigualdad propuesta. En efecto, la teoría $T$ de Newton podría ser considerada como establecida firmemente en 1846, aun antes del descubrimiento de Neptuno, y, así, podría parecer absurdo atribuir a $T$ una credibilidad tan baja. No creo que estemos obligados a considerar una credibilidad $10^{-5}$ tan baja en este caso: podríamos pensar que la certeza lógica, a la que adscribimos la credibilidad 1, es incomparablemente mayor que la confianza que colocamos en la mejor generalización inductiva establecida, cuya credibilidad podríamos considerar aún como infinitesimal; cf. ejemplo 8.

Después de todo esto podemos encontrar más seguro volver al punto de partida de la sección 5-10, al cual nos adherimos esencialmente en (4). Representémonos cualitativamente cómo un cambio de este o aquel componente de la situación influiría en nuestra confianza, pero no nos entreguemos a ningún estimado cuantitativo.

**5.** *Verosimilitud y credibilidad.* Podemos tener una conjetura sobre probabilidades. Por ejemplo, que el dado que tenemos en nuestras manos es idealmente «justo», es decir, que cada cara tiene la misma probabilidad 1/6. Naturalmente, esta conjetura es dura y difícil de creer. O podemos conjeturar que cada cara del dado tiene una probabilidad entre 0,16 y 0,17, lo cual es ya más creíble. Una conjetura sobre probabilidades es una hipótesis estadística. A menudo ocurre que tenemos sólo dos conjeturas rivales obvias: una conjetura física $F$ y una hipótesis estadística $H$; cf. sección 14.9 (7) y ejemplo 14.33. En tal caso, podemos referirnos seriamente a $\Pr\{H\}$, la credibilidad de la hipótesis estadística $H$.

Tal hipótesis se prueba apropiadamente por observación estadística. Supongamos que $E$ denomina la predicción de que la observación estadística arrojará tal y tal resultado. Consideremos la credibilidad $\Pr\{E/H\}$ y supongamos que esta credibilidad tiene un valor numérico, igual a la probabilidad de que un acontecimiento de la clase predicha por él ocurrirá, calculado sobre la base de la hipótesis estadística $H$. Como he-

mos visto en el ejemplo 4, ciertas suposiciones (completamente naturales) de intercambiabilidad o simetría (incluidas en la hipótesis estadística $H$) pueden incluso obligarnos a igualar credibilidad con probabilidad.

Esta credibilidad o probabilidad, $\Pr\{E/H\}$, puede ser considerada desde dos puntos de vista diferentes, como hemos discutido en la sección 14.7 (5); cf. también la sección 14.8 (5). Por otro lado, $\Pr\{E/H\}$ es la probabilidad de un suceso de la clase predicha por $E$, calculada sobre la base de la hipótesis estadística $H$. Por otro, si tal acontecimiento ocurre realmente y es observado, nos inclinamos a pensar que $H$ es menos verosímil cuanto menor sea el valor numérico de $\Pr\{E/H\}$, y por esta razón llamos a $\Pr\{E/H\}$ la *verosimilitud* de la hipótesis estadística $H$, juzgada a la vista del hecho de que el suceso predicho por $E$ ha ocurrido realmente. Cf. sección 14.7 (5) y también sección 14.8 (5).

Ahora, se sigue del ejemplo 14.26 (2) que

$$\Pr\{H/E\} = \frac{\Pr\{E/H\}\,\Pr\{H\}}{\Pr\{E\}}.$$

En esta ecuación, $\Pr\{E/H\}$ no es sólo una credibilidad, sino también una probabilidad, y tiene un valor numérico definido. No obstante, $\Pr\{H/E\}$, $\Pr\{H\}$ y $\Pr\{E\}$ son sólo credibilidades y no suponemos que posean valores numéricos definidos*. Especialmente, $\Pr\{H\}$ y $\Pr\{H/E\}$ denotan la credibilidad de la misma hipótesis estadística $H$, pero la primera antes, y la segunda después, de la observación del suceso predicho por $E$. Volvamos a formular la ecuación anterior en una forma menos convencional, haciendo hincapié en uno de los aspectos de $\Pr\{E/H\}$:

$$\text{Credibilidad después del suceso} = \frac{\text{Verosimilitud} \times \text{Credibilidad antes del suceso}}{\text{Credibilidad del suceso}}.$$

Habiendo observado el suceso predicho por $E$ nos enfrentamos con una decisión: ¿rechazaremos la hipótesis estadística $H$ y aceptaremos la conjetura rival física $F$, o qué haremos? Nuestra decisión se basará en la última información, y, por tanto, en $\Pr\{H/E\}$, la credibilidad de la hipótesis estadística después de la observación del suceso. La verosimilitud $\Pr\{E/H\}$ es un factor de la credibilidad $\Pr\{H/E\}$: es el más importante, quizá, porque posee un valor numérico, computable por un claro familiar procedimiento, pero es aún sólo un factor, no la expresión plena de la credibilidad.

* Esta es la situación normal. Sólo excepcionalmente está el estadístico en posición de adscribir un valor numérico a $\Pr\{H\}$.

La verosimilitud es una indicación importante, pero no todo. El estadístico puede limitarse sabiamente al cálculo de la verosimilitud, pero el cliente del estadístico puede actuar insensatamente si desdeña los otros factores. Este pesará cuidadosamente $\Pr\{H\}$, la credibilidad de la hipótesis estadística $H$ antes del suceso: queremos decir, en efecto, esta $\Pr\{H\}$ cuando hablamos de lo «apropiado», o del «realismo» de $H$. Cf. ejemplo 14.33.

**6.** *El intento de Laplace para enlazar inducción con probabilidad.* Un saco contiene bolas blancas y negras en proporción desconocida; $m$ blancas y $n$ negras han sido sucesivamente sacadas y reemplazadas. ¿Cuál es la probabilidad de que $m' + n'$ subsiguientes extracciones den $m'$ bolas blancas y $n'$ negras?

Un caso particular de este problema aparentemente sencillo sobre el saco y las bolas puede ser interpretado como el problema fundamental sobre la inferencia plausible reducido a su expresión más sencilla. En efecto, consideremos el caso en que $n = n' = 0$. Sacamos $m$ bolas del saco de composición desconocida, y todas las bolas resultan ser blancas. Podemos asimilar esta situación a la del naturalista que comprueba $m$ consecuencias de una conjetura y encuentra todas las $m$ observaciones concordantes con esa conjetura. El naturalista proyecta más observaciones. ¿Cuál es la probabilidad de que su próxima observación $m'$ resulte concordante con la conjetura? Esta pregunta puede ser ideada como el caso particular $n = n' = 0$ del problema propuesto sobre el saco y las bolas.

En este problema hay un punto oscuro y confuso: la proporción de las bolas negras y las blancas del saco es desconocida. Sin embargo, en vista de la interpretación propuesta, este punto aparece como esencial: el naturalista no puede conocer el «funcionamiento interno» de la naturaleza. El conoce sólo lo que ha observado, y nosotros sabemos sólo lo que el saco nos da: hasta este momento, tantas y tantas bolas blancas en tantas y tantas extracciones.

El sentido común nos sugiere que no podemos calcular la probabilidad requerida si no sabemos nada de la composición del saco; un problema sin datos suficientes es insoluble. Sin embargo, Laplace dio una solución —una solución controvertible—. ¿Cómo se las arregló para llegar a una solución?

Laplace introdujo un nuevo principio para compensar la falta de datos; este principio es controvertible. «Cuando la probabilidad de un simple suceso es desconocida podemos suponer todos los valores de esta probabilidad entre 0 y 1 como igualmente verosímiles», dice

Laplace *. «Esto es la distribución de la ignorancia por igual», se mofan sus oponentes.

Una vez que el principio controvertible de Laplace es admitido, la derivación viene en línea recta hacia nosotros; no necesitamos considerarla aquí. Su resultado es: si $m$ extracciones dan sólo bolas blancas, la probabilidad de que $m'$ subsiguientes extracciones den también sólo bolas blancas es

$$\frac{m+1}{m+m'+1}.$$

Llamemos a este enunciado la «regla general de sucesión». El mejor caso particular conocido se refiere a $m' = 1$: si $m$ extracciones dan sólo bolas blancas, la probabilidad de que también la próxima extracción dé una bola blanca es

$$\frac{m+1}{m+2}.$$

Llamemos a ésta la «regla particular de sucesión»**.

¿Son aceptables estas reglas si interpretamos «bolas blancas» como «observaciones concordantes de la misma naturaleza» y «probabilidad» como «grado de confianza razonable»? Esta pregunta da en el clavo y la discutiremos.

(1) Volvamos a considerar nuestro primer ejemplo de razonamiento inductivo. La conjetura de Goldbach asegura que, desde $6 = 3 + 3$ en adelante, cualquier entero par es la suma de dos primos impares. La tabla de la sección 1.3 verifica esta conjetura hasta 30. Habiéndola verificado hasta 30 esperamos, con más o menos confianza, que será también verificada en el caso próximo, 32. La regla particular de sucesión puede interpretarse de modo que signifique que, habiendo verificado la conjetura de Goldbach en los primeros $m$ casos, estamos autorizados a esperar su verificación en el caso próximo con la probabilidad

$$\frac{m+1}{m+2} = 1 - \frac{1}{m+2}.$$

Démonos cuenta de lo que esto significa. A medida que $m$ aumenta, la probabilidad también aumenta: en efecto, cuanto más casos han sido

* *Oeuvres complètes*, vol. 7, págs. XVII.

** Esto difiere algo de la terminología usual, Cf. J. M. KEYNES, *A treatise on probability*, páginas 372-383.

verificados en el pasado, más confiadamente esperamos que la conjetura se verifique en el caso próximo. Si $m$ tiende a $\infty$, la probabilidad tiende a 1: podríamos esperar aproximarnos más y más, recogiendo más y más verificaciones. Consideremos ahora la diferencia de dos probabilidades, una que corresponde a $m + 1$, y la otra a $m$, verificaciones previas:

$$\frac{m+2}{m+3} - \frac{m+1}{m+2} = \frac{1}{(m+2)(m+3)}.$$

Esta diferencia disminuye cuando $m$ aumenta: es cierto que cada nueva verificación aumenta nuestra confianza, pero la aumenta cada vez menos cuando viene después de más y más verificaciones previas semejantes. (La semejanza de las verificaciones es esencial en este punto; cf. sección 12.2.)

Tomemos ahora la regla general de sucesión. Esta podría ser interpretada como significado que: habiendo verificado la conjetura de Goldbach en los primeros $m$ casos, estamos autorizados a esperar su verificación en los próximos $m'$ casos con la probabilidad

$$\frac{m+1}{m+m'+1}.$$

Si mantenemos a $m$ fijo, aunque $m'$ aumente, esta probabilidad disminuye: en efecto, cuanto más intentemos predecir el futuro sobre la base de la observación pasada menos confiadamente podremos predecirlo. Si $m'$ aumenta indefinidamente, la probabilidad tiende a 0. En efecto, la verificación para todos los valores de $m'$ significará que la conjetura de Goldbach es cierta. Obviamente, sobre la base de un número dado $m$ de observaciones no podemos asegurar que la conjetura sea cierta. La regla parece implicar un enunciado más fuerte: incluso sobre la base de $m$ observaciones no podemos atribuir una probabilidad diferente de 0 a la conjetura de Goldbach. Y un enunciado tan fuerte puede apuntar en la dirección correcta.

(2) Hasta este punto la regla de sucesión parece bastante respetable. Sin embargo, mirémosla más concretamente. Atribuyamos valores numéricos a $m$ y no desdeñemos las situaciones de la vida diaria. Esto será suficiente para considerar la regla particular de sucesión.

He comprobado los números pares 6, 8, 10, ... 24, y encuentro que cada uno de ellos es una suma de dos primos impares. La regla dice que debo esperar con la probabilidad 11/12 que 26 sea también la suma de dos primos impares.

En una ciudad extranjera, donde escasamente puedo entender la lengua, he comido en un restaurante donde se produjeron graves malentendidos. Sin embargo, después de diez comidas no noto los efectos de ninguna enfermedad y me dirijo confiadamente al restaurante la undécima vez. La regla dijo que la probabilidad de que no fuese envenenado en mi próxima comida era de 11/12.

Un muchacho cumple hoy los diez años. La regla dice que habiendo vivido diez años tiene la probabilidad 11/12 de vivir uno más. El abuelo del muchacho llegó a los setenta años. La regla dice que tiene la probabilidad 71/72 de vivir un año más.

Estas aplicaciones parecen tontas, pero ninguna más tonta que la siguiente debida al mismo Laplace. «Suponemos —dice— que la historia lleva cinco mil años, es decir, 1.826.213 días. El Sol sale cada día, y, así, podemos apostar 1.826.214 contra 1 que el Sol saldrá mañana de nuevo» *. Me cuidaría mucho, desde luego, de no presentar una apuesta semejante a un colega noruego que nos organizase un vuelo a algún lugar del Círculo Artico.

Pero la regla puede incluso con este absurdo. Apliquémosla al caso $m = 0$: la derivación de la regla es tan válida para este caso como para cualquier otro. Sin embargo, para $m = 0$ la regla asegura que cualquier conjetura sin ninguna verificación tiene la probabilidad 1/2. Cualquiera puede inventar ejemplos para mostrar que tal aserción es monstruosa. (A propósito, ésta es también autocontradictoria.)

(3) Nuestra discusión ha sido larga. En lenguaje más cauto habría sido aún más larga y podría haberse prolongado, pero el resultado es que: la regla de sucesión puede parecer sabia si evitamos los valores numéricos, pero resulta ser, ciertamente, una locura si le atribuimos tales valores. Quizá esto apunta a una enseñanza: aplicar el cálculo de probabilidad al razonamiento plausible y evitar los valores numéricos en principio. En todo caso, éste es el punto de vista por el que abogamos en el presente capítulo.

**7.** *¿Por qué no cuantitativo?* Este capítulo presenta una tesis: El cálculo de probabilidad deberá ser aplicado al razonamiento plausible, pero sólo cualitativamente. Existe, sin embargo, una fuerte tentación de aplicarlo cuantitativamente, y así, tenemos que examinar unos cuantos puntos notables.

(1) *No comparable.* Hay cierta evidencia de que la conjetura de Goldbach, que se refiere a la suma de dos primos impares, es correcta, ver secciones 1.2-1.3. Hay alguna evidencia de que los nórdicos llegaron

* *Oeuvres complètes*, vol. 7, pág. XVII.

a América algunos cientos de años antes que Colón. ¿Qué evidencia es más fuerte?

Esto parece ser una pregunta muy tonta, realmente. ¿Cuál podría ser el propósito de comparar dos casos tan dispares? Y ¿quién los compararía? Para juzgar la evidencia competentemente habría que ser un experto. En un caso, la evidencia debe ser juzgada por un matemático experto en la teoría de los números. En el otro caso, la evidencia debe ser juzgada por un historiador, experto en islandés antiguo. Habrá pocas personas que sean expertas en ambas materias. No obstante, puede ocurrir que nuestra aparentemente tonta pregunta apunte a una posibilidad. Podría ocurrir que no hubiese decisión ni manera razonable de decir cuál evidencia es la más fuerte. Esta posibilidad es tan importante que merece un nombre. Si no hay modo razonable de decidir qué evidencia es más fuerte, $S_1$ o $S_2$, llamaremos a $S_1$ *no comparable* con $S_2$. Podríamos encontrar en los anteriores capítulos varios ejemplos, que sugieren más clara y convincentemente, que el ejemplo con que empezamos aquí, que una evidencia no sea comparable con otra. Ver sección 4.8.

(2) *Comparable.* Después de haber examinado la posibilidad de evidencias no comparables, examinemos ahora varios casos de evidencias obviamente comparables.

$S_1$ denota la evidencia de la conjetura de Goldbach (sección 1.2), que resulta de su verificación para todos los números pares hasta 1.000. $S_2$ denota la evidencia para la misma conjetura, que resulta de su verificación hasta 2.000. Obviamente, la evidencia $S_2$ es más fuerte que la evidencia $S_1$.

Ahora, cambiemos la notación, y $S_1$ denota la evidencia presente para la llegada al continente americano de nórdicos antes que Colón. $S_2$ denota lo que esta evidencia podría aumentar, si alguien descubrió, digamos, un cementerio en cualquier parte de las costa del Labrador con escudos y espadas semejantes a los de la edad de los vikingos. Obviamente, la evidencia $S_2$ sería más fuerte que la evidencia $S_1$.

Consideremos un caso más sutil. Ahora $s_1$ denota la evidencia de una conjetura $A$, que resulta de la verificación de una de sus consecuencias $B$. Después que la consecuencia $B$ ha sido verificada, alguien observa que $B$ es muy improbable en sí misma. (Esta observación podría ser bastante precisa; la probabilidad de $B$, calculada sobre las bases de una sencilla y aparentemente adecuada hipótesis estadística, podría ser baja.) Esta observación cambia la evidencia de la conjetura $A$ en $s_2$. La evidencia $S_2$ es más fuerte que la evidencia $S_1$. (Hemos dicho y aesto, quizá algo menos agudamente, en la sección 12.3.)

En los tres ejemplos hemos obtenido la evidencia $S_2$ de la evidencia $S_1$

añadiendo algunas observaciones relevantes. Pero, si no existe tan sencilla relación entre $S_1$ y $S_2$ ¿cómo podríamos decidir cuál es más fuerte? Esta pregunta nos acerca a la posibilidad de que la evidencia sea no comparable.

(3) *Comparable, pero aún no cuantitativo.* En la anterior subsección (2) hemos visto casos en que una evidencia $S_2$ puede ser más fuerte que una evidencia $S_1$. Pero *¿cuánto más fuerte?* Me parece que no hay contestación razonable a esta pregunta en los casos anteriores. Y, así, vemos que todavía estamos en un nivel cualitativo.

(4) *¿Cómo parecerá?* En la sección 4 tomamos el símbolo $\Pr\{A\}$ para denotar la credibilidad de una conjetura $A$. En las secciones siguientes de este capítulo intentamos seguir sin dar un determinado valor numérico a $\Pr\{A\}$: en esto consiste el punto de vista «cualitativo» por el que este libro aboga. El punto de vista «cuantitativo» consistirá en dar a $\Pr\{A\}$ un valor numérico definido siempre que el cálculo de probabilidad sea aplicado al razonamiento plausible en la conjetura $A$.

La prueba debe correr a cargo de los que defienden una aplicación del cálculo de probabilidad al razonamiento plausible. Todo lo que éstos deben hacer es ingeniarse una clase de conjeturas $A$ no triviales, cuya credibilidad $\Pr\{A\}$ pueda ser calculada por métodos que conducen a resultados aceptables, al menos en algunos casos *.

Nadie ha propuesto todavía un método claro y convincente para calcular credibilidad en casos no triviales, y, si hemos visualizado situaciones concretas en que un estimado de probabilidades es importante (como hemos hecho), podemos percibir fácilmente que cualquiera atribución de valores numéricos definidos a las credibilidades es un gran peligro.

Las probabilidades que tienen valores numéricos definidos son comparables: dos números, o son iguales, o uno es más grande que otro. Sin embargo, tras las discusiones de (1) y (2) encontramos difícil de aceptar que dos conjeturas cualesquiera son comparables en credibilidad. Tomemos las dos conjeturas con que empezamos nuestra discusión: la conjetura de Goldbach y la conjetura histórica sobre el descubrimiento de América. Si atribuimos valores numéricos a sus credibilidades, la fuerza de la evidencia que habla por una podría ser comparada con la de la otra, aunque tal comparación parecerá fútil y absurda.

(5) *¿Merecerá la pena?* Hay otro punto a considerar que es independiente de la discusión anterior. El peso de un argumento plausible

* Atribuir valor numérico a una credibilidad sobre la base de una asunción de intercambiabilidad o simetría (como hicimos en el ejemplo 4) sería muy trivial: se requiere algo más, y más nuevo, para justificar credibilidades cuantitativas.

puede ser extremadamente importante, pero tal importancia es provisonal, efímera, pasajera: ¿Valdrá la pena establecer un valor numérico sobre algo tan transitorio?

¿Cuál es la credibilidad de la ley de la gravitación de Newton, juzgada a la luz de los hechos recogidos en la primera edición de los *Principia?* Imaginemos por un momento que existe un método para evaluar numéricamente tal credibilidad. Sin embargo, no imaginemos que la evaluación pueda ser fácil: en vista de la complejidad de los hechos y sus interrelaciones, la evaluación debe ser delicada y el cálculo numérico largo. ¿Merece la pena emprenderlo? Quizá para nosotros, teniendo en cuenta la importancia histórica y filosófica del descubrimiento de Newton. Pero escasamente para Newton y sus contemporáneos: en lugar de calcular la credibilidad de la teoría, ellos podrían, con el mismo esfuerzo, haber cambiado la credibilidad por el desarrollo de la teoría, y multiplicar las observaciones. Parece descabellado dedicar diez años al cálculo de un grado de credibilidad que durará sólo un segundo.

**8.** *¿Credibilidades infinitesimales?* Un nuevo caso particular de alguna conjetura numérico-teórica (tal como la conjetura de Goldbach; véanse secciones 1.2-1.3) ha sido verificado. Tal verificación debe considerarse que aumenta el peso de la evidencia, o la credibilidad de la conjetura. Pero ninguna cantidad de tales verificaciones puede probar la conjetura. Incluso podemos sentir que ninguna cantidad de verificaciones semejantes traerá la conjetura ostensiblemente próxima a una prueba. (Cf. ejemplos 4 (5) y 6 (1).) Tales sentimientos pueden sugerir la introducción de infinitesimales en el cálculo de probabilidad.

Los infinitesimales pueden ser manejados con toda claridad por los matemáticos modernos. Consideramos «cantidades» $a$, $b$, ... representadas por las «series formales de valores»:

$$a = a_0 + a_1\varepsilon + a_2\varepsilon^2 + a_1\,\varepsilon^3 + \ldots;$$

$a_0, a_1, a_2, \ldots$ son números reales, $\varepsilon$ es un número indeterminado y no se presta atención a la convergencia. Hay un álgebra de tales cantidades; hay reglas (familiares a la teoría de las series convergentes de valores) según las cuales tales series formales de valores pueden ser sumadas, restadas y multiplicadas; incluso la división por $a$ puede realizarse si $a_0 \neq 0$. Llamamos $a$ la «cantidad cero» o cero cuando $a_0, a_1, a_2, \ldots$ desaparecen. Dos cantidades son iguales si su diferencia es 0. La cantidad $a$ se reduce al número $a_1$ cuando $a_1, a_2, a_3, \ldots$ desaparecen.

Decimos que $a$ es *positivo* si el *primer* número, que no desaparece en la secuencia $a_0, a_1, a_2, \ldots$ es positivo. De esta definición podemos derivar fácilmente dos reglas básicas:

O $a$ es 0, o $a$ es positivo, o $-a$ es positivo, y estas tres posibilidades son mutuamente excluyentes.

La suma y el producto de dos cantidades positivas son positivos.

Decimos que $a > b$ si $a - b$ es positivo. De estas definiciones se sigue que cualquier *número* positivo es mayor que $\varepsilon$ (es decir, la serie formal de valores $0 + \varepsilon + 0 \cdot \varepsilon^2 + 0 \cdot \varepsilon^3 + \ldots$, aunque $\varepsilon$ sea positivo. Y, así, es un infinitesimal «real».

Podemos contemplar ahora alguna forma del cálculo de probabilidad en que las probabilidades (o credibilidades) no son números necesariamente, sino cantidades $a$ de la clase introducida, sujetas a la condición $0 \leqq a \leqq 1$. En tal cálculo de credibilidad de la conjetura de Goldbach, después de 1.000.000, o cualquier otro número de verificaciones, podría aún ser infinitesimal, es decir, una cantidad $a$ con $a_0 = 0$.

Me abstengo de comentarios sobre las perspectivas de un plan semejante. En cualquier caso, manifiesta una posibilidad de examinar credibilidades no numéricas desde otro ángulo.

**9.** *Reglas de admisibilidad.* Mi deseo es que el lector saque su propia conclusión y se forme su opinión. En consecuencia, he pospuesto mi opinión sobre una cuestión crucial hasta este último comentario del capítulo. Me refiero a la pregunta sugerida por la primera sección de éste: ¿Tiene el razonamiento plausible reglas de alguna clase?

Es bastante obvio que el razonamiento plausible no tiene reglas de la misma clase que el demostrativo. Se ha propuesto una demostración. Si ésta se presenta por medio de pasos lo suficientemente pequeños, la validez de cada paso será comprobada por una regla de lógica formal. Si todos los pasos se adecuan a las reglas de demostración, es válida, pero será inválida si uno de estos pasos viola las reglas. Así, las reglas de la lógica demostrativa son decisivas: pueden decidir si un argumento demostrativo propuesto está bien trabado o no. Los patrones de razonamiento plausible no pueden lograr nada semejante. Un argumento plausible ha sido propuesto. Cada paso de éste trata de hacer más digna de crédito a cierta conjetura, siguiendo algún patrón aceptado. Habiendo seguido el argumento paso a paso, usted no está sujeto a confiar en la conjetura, en ningún grado determinado.

Sin embargo, hay diferentes clases de reglas. Las reglas lógicas son muy diferentes de las reglas legales. Un tribunal judicial escuchará a todas las partes concernientes, pero no tendrá en cuenta las irrelevancias. Por tanto, dicho tribunal tendrá poder para excluir las cuestiones irrelevantes de sus procedimientos, y tal poder estará regulado por *reglas de admisibilidad.* Sin alguna regla de esta clase no podría haber admi-

nistración de justicia: el tribunal no podría rechazar a un abogado sin escrúpulos que, por cuestiones irrelevantes o indebidas, presentara testimonios adversos a la acusación, al jurado y al juez, o alargase los procesos indefinidamente.

Los patrones de razonamiento plausible, recogidos en lo anterior, pueden ser considerados como *reglas de admisibilidad en la discusión científica.* No estamos obligados en absoluto a dar un determinado grado de creencia a una conjetura, si algunas de sus consecuencias han sido verificadas. Pero si la conjetura es discutida no cabe duda que es admisible mencionar tales verificalones y resulta justo y razonable escucharlas. Nuestros patrones registran varios puntos que conciernen a tales verificaciones, influyentes en el peso de la evidencia (como la analogía, o la falta de analogía, con verificaciones anteriores, etc.). Es justo y razonable admitir la discusión de tales puntos. Al recoger estos patrones la intención del autor no fue otra que enumerar estos puntos generales que, *de acuerdo con los usos de los buenos científicos, son admisibles en una discusión científica,* con un punto de vista que influye en la credibilidad de la conjetura examinada.

En un juicio los poderes del tribunal están divididos entre el jurado y el presidente. Esta división de poderes (como concebida por ciertas autoridades legales y aceptada con alguna extensión por la práctica judicial de ciertos Estados y países) es de gran interés para nosotros. El jurado y el juez tienen funciones diferentes, deciden diferentes cuestiones. Las cuestiones que se refieren a la admisibilidad de evidencias son contestadas por el juez, las que conciernen a la credibilidad de la evidencia admitida son contestadas por el jurado. El juez ha de decidir qué evidencia merece la consideración del jurado. Este debe determinar si la evidencia sometida es de suficiente peso. Al decidir qué cuestiones y cuáles no son de la consideración del jurado, el juez debe conocer y respetar los precedentes, los usos del tribunal y las reglas de admisibilidad establecidas. Al pesar la evidencia sometida, el jurado, que posiblemente carece de instrucción legal, debe confiar en sus propias luces naturales.

En resumen: los poderes del tribunal están divididos entre el juez y el jurado, como el poder de juzgar una conjetura propuesta está dividido, en cada uno de nosotros, entre las reglas impersonales y el buen sentido personal. El juez representa el papel de la regla; el jurado, el de nuestro discernimiento personal. Las reglas impersonales de razonamiento plausible están para decidir qué clase de evidencia merece consideración. Mientras que el buen sentido personal está para decidir si el fragmento particular de evidencia sometido tiene peso suficiente o no.

# El razonamiento plausible en la invención y en la enseñanza

*Las palabras constan de letras del alfabeto; las frases, de palabras que pueden encontrarse en el diccionario, y los libros, de frases que pueden también encontrarse en otros autores. Pero si las cosas que digo son consistentes y están conectadas de modo que se sigan unas de otras, usted puede inculparme de haber cogido mis frases de otros, como de haber cogido mis palabras del diccionario.*—DESCARTES *.

## 1. OBJETO DEL PRESENTE CAPÍTULO

Los ejemplos de la primera parte de esta obra y las discusiones de los anteriores capítulos de la segunda parte espero que hayan elucidado un poco el papel del razonamiento plausible en el descubrimiento de los hechos matemáticos. Pues el matemático no sólo intuye; también resuelve problemas y tiene que probar los hechos que ha intuido. ¿Cuál es el papel del razonamiento plausible en el descubrimiento de la solución o en la invención de la prueba? Esta es la pregunta que será contestada en el presente capítulo. Y, por cierto, ésta es la pregunta que atrajo al autor, quien, primeramente envuelto en los métodos de resolver problemas, fue poco a poco conducido al tema del presente libro.

El tema del razonamiento plausible es sutil y elusivo, y lo mismo ocurre con los métodos de solución. Por ello quizá fue adecuado diferir la cuestión que combina dos temas tan delicados hasta el último capítulo. La atención que le prestaremos será breve; el principal propósito es apuntar a la conexión con materias previamente discutidas. Un tratamiento más amplio sería más propio de otro libro sobre métodos de resolver problemas.

## 2. LA HISTORIA DE UN PEQUEÑO DESCUBRIMIENTO

La solución de un sencillo problema matemático, aunque no meramente rutinario, puede traernos la tensión y el triunfo de un descubri-

* *Oeuvres*, editadas por Adam y Tannery, vol. 10, 1908, pág. 204.

miento. Miremos el siguiente ejemplo: *Construir un cuadrilátero, siendo dados sus cuatro lados, a, b, c,* y *d, y* ε *el ángulo formado por los lados opuestos a y c.*

Los datos del problema se muestran en la figura 16.1: cuatro líneas y un ángulo, fragmentos de una figura hecha pedazos, que reuniremos para satisfacer todos los requisitos establecidos en el problema.

Se enticnde que los lados *a, b, c* y *d* se siguen uno al otro en este orden en torno al deseado cuadrilátero, de manera que *a* es opuesto a *c,* y *b* a *d.* El ángulo ε, formado por los lados opuestos *a* y *c, no* es uno de los cuatro ángulos del cuadrilátero.

Hagamos algunas de las preguntas usuales que pueden acercarnos al problema.

*¿Son suficientes los datos para determinar la incógnita?* Los cuatro lados serían insuficientes, como es obvio, para determinar el cuadrilátero: cuatro varillas unidas por medio de articulaciones flexibles en sus extremos respectivos forman un cuadrilátero articulado, móvil, deformable, flexible y sin forma determinada. Pero si fijamos uno de sus cuatro ángulos el cuadrilátero no puede moverse más: el cuadrilátero queda determinado por sus cuatro lados y uno de sus ángulos. Podemos adivinar que está también determinado por cuatro lados y algún otro ángulo, y, así, los datos de nuestro problema parecen suficientes.

*Dibujar una figura.* Dibujamos la figura 16.2, que desarrolla los cinco datos reunidos de acuerdo con las condiciones del problema propuesto. Seguramente tendremos que *usar todos los datos.*

Puede ocurrir que nos adhiramos a este punto y no se nos ocurra ninguna idea útil durante un rato. En efecto, la figura 16.2 parece tosca. Los lados *a, b, c, d* están en sus propios lugares, desde luego, pero la situación del *ángulo* ε parece desafortunada. Este ángulo es uno de nuestros datos y tendremos que usarlo. Sin embargo, ¿cómo podremos usarlo estando situado tan lejos y en un lugar tan poco corriente?

Un experto en resolver problemas intentaría *volver a dibujar* la figura: intentaría colocar ese ángulo ε en alguna otra parte. Puede acertar así en la figura 16.3, donde el ángulo ε está entre el lado *a* y una paralela al lado *c,* trazado por el punto medio de *a.* La figura 16.3 parece más prometedora que la figura 16.2.

¿Por qué parece la figura 16.3 tan prometedora? Aun buenos estudiantes, que están muy seguros de este hecho, pueden ser incapaces de contestar a la pregunta claramente:

«Me parece bastante buena.»

«Los datos están conjuntamente adecuados.»

Sólo un estudiante de talento excepcional o muy experimentado

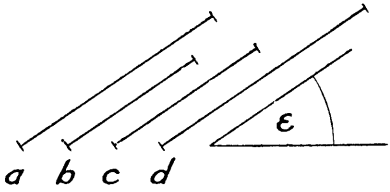



Fig. 16.1. Fragmentos.

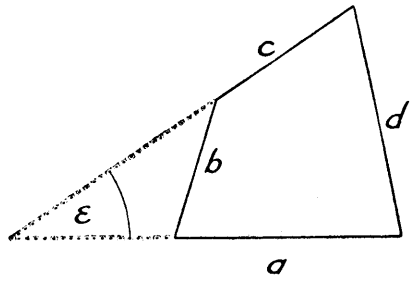



Fig. 16.2. Obvio.

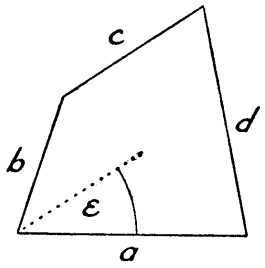



Fig. 16.3. Tibio.

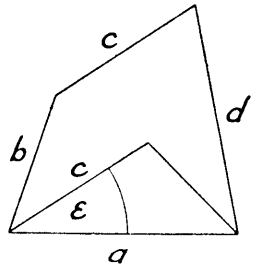



Fig. 16.4. ¡Caliente!

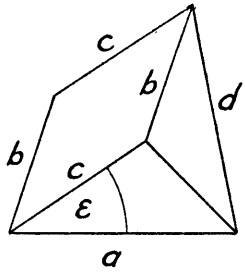



Fig. 16.5. ¡Hecho!

será capaz de dar una explicación completa: «En la figura 16.2 el ángulo está situado en un triángulo. Sin embargo, este triángulo no es adecuado para la construcción: sólo son conocidos dos datos, ε y *b*. Según está situado en la figura 16.3, el ángulo ε tiene más oportunidad de servirnos para un triángulo adecuado. Esto es deseable, pues, normalmente, esta clase de construcción se reduce a la construcción de un triángulo con los datos adecuados»

La idea general, que hay tras esta última contestación, parece ser ésta: cualquier rasgo en que la situación presente se halla de acuerdo con afortunadas situaciones pasadas parece prometedor.

Sea lo que sea, la figura 16.3 está dentro de nuestras expectaciones. En efecto, el ángulo ε se adapta a un triángulo, que podemos construir fácilmente (a partir de los lados *a* y *c* y el ángulo comprendido entre ellos ε; ver figura 16.4). Cuando este triángulo ha sido terminado la solución está bastante cerca. De hecho, al triángulo terminado se une otro, que podemos construir (a partir de los lados *b*, *d* y un lado previamente construido; ver figura 16.5). Una vez dibujados ambos triángulos completamos la construcción buscada, mediante el trazo del paralelogramo de lados *b* y *c*.

Para la mayoría de los que trabajan en problemas, el paso decisivo en la solución anterior, representado por una asociación de cinco figuras, es el paso de la figura 16.2 a la figura 16.3; un largo período de duda puede preceder a este paso. Pero, una vez que el ángulo ε es colocado convenientemente, el progreso de la solución, desde la figura 16.3 hasta la figura final 16.5, pasando por la figura 16.4, puede ser bastante rápido.

La anterior solución puede revelar uno o dos puntos, que juegan también un papel en descubrimientos más importantes.

## 3. El proceso de solución

Resolver un problema es un proceso extremadamente complejo. Ninguna descripción o teoría de ese proceso puede agotar sus múltiples aspectos; cualquier descripción o teoría del mismo está abocada a ser incompleta, esquemática y muy simplificada. Mi deseo es señalar el lugar del razonamiento plausible en este complejo proceso y elegiré la descripción más sencilla que pueda encontrar para que dicho lugar sea visiblemente localizado, Y, así, el mero principio de una descripción tan sencilla nos será suficiente.

(1) *Proponiéndonos un problema a nosotros mismos.* Un problema se hace problema para nosotros cuando nos lo proponemos como tal.

No es todavía nuestro porque tengamos que resolverlo en un examen. Si deseamos que alguien venga y nos dé la respuesta sospecho que todavía no nos hemos propuesto el problema con seriedad a nosotros mismos. Pero si estamos ansiosos de encontrar la contestación por nuestra cuenta, con nuestros propios medios, entonces hemos hecho realmente nuestro el problema, estamos seriamente interesados en él.

Proponernos el problema a nosotros mismos es el principio de la solución, el movimiento primero y esencial en el juego. Es un movimiento que tiene la naturaleza de una decisión.

(2) *Atención selectiva.* No necesita usted decirme que se ha propuesto el problema a sí mismo, necesita decírselo a usted mismo: su conducta total mostrará que lo hizo. Su mente se hace selectiva; se hace más sensible a cualquier cosa que parezca estar conectada con el problema y menos sensible a lo que no parezca estar conectado con él. Usted recoge ávidamente cualquier observación, sugestión o hecho que pueda ayudarle a resolver el problema y cierra la puerta a otras cosas. Cuando la puerta está tan herméticamente cerrada que incluso las más urgentes llamadas del mundo externo dejan de afectarle, la gente dice que usted está absorto.

(3) *Registrando la marcha del progreso.* Hay todavía otra cosa que muestra que usted está seriamente comprometido en su problema; usted se hace más sensible. Siente agudamente la marcha de su progreso; se siente animado si éste es rápido, deprimido si es lento. Todo lo que llega a su mente es clasificado: «Parece bueno», «Podrían ayudarme», o «No es bueno», «No me ayuda». Tales juicios no son, por supuesto, infalibles. (Aunque parecen ser a menudo más correctos que otra cosa, especialmente con gente de talento o experimentada.) En cualquier caso, tales juicios y sentimientos son importantes para usted personalmente; ellos guían su esfuerzo.

(4) *Dónde empieza el razonamiento plausible.* Veamos, algo más concretamente, una situación típica.

Usted intenta alcanzar la solución en cierta dirección y siguiendo cierta línea. (Por ejemplo, tratando de resolver el problema geométrico de la sección 2, usted rechaza la figura 16.2 e intenta trabajar con la figura 16.3, que le parece más esperanzadora.) Usted puede sentir agudamente que trabaja en la dirección correcta, que sigue una prometedora línea de enfoque, que está usted en la pista. Puede sentir así, desde luego, sin formular sus sentimientos en palabras. E incluso si dice algo como «Parece bueno» no se toma la molestia de analizar su confianza, no pregunta «¿Por qué parece bueno?». Está demasiado ocupado siguiendo la pista.

Pero puede usted tener mala suerte. Tropieza con dificultades, no hace muchos progresos, nada nuevo le ocurre y empieza a dudar: «¿Fue un buen comienzo? ¿Es ésta la dirección correcta?» Y entonces puede empezar a analizar sus sentimientos. «La dirección parecía totalmente plausible —pero ¿por qué es plausible?». Comienza a debatirse consigo mismo y se le ocurren algunas razones más claras:

«La situación no es tan mala. Podría construir un triángulo. Siempre se empieza construyendo triángulos en estos problemas».

«Seguramente, éste fue un comienzo correcto, después de todo. Esta parece la solución correcta. ¿Qué necesito para solucionar esta clase de problemas? Tal punto —y yo lo tengo—. Y esa clase de puntos —también lo tengo—. Y...».

Sería interesante ver cómo la gente razona en una situación semejante —en efecto, éste es nuestro principal propósito—. Pero aún necesitamos un ejemplo para ampliar nuestra base de observaciones.

## 4. «Deus ex machina.»*

El próximo ejemplo será un poco menos sencillo que el de la sección 2. Este ejemplo lo daremos en la sección 6, tras algunos preparativos en ésta y en la próxima. La sección 6 mostrará una prueba, presentada de un modo que contrasta con la manera normal de presentación. Para hacer hincapié en el contraste veamos primero la prueba como sería presentada en un avanzado libro de texto o en una revista matemática.

Un libro o lectura matemática debe ser, ante todo, correcto y sin ambigüedades. Sabemos, por dolorosa experiencia, que una exposición correcta y sin ambigüedades puede estar lejos de ser satisfactoria y puede parecer falta de inspiración, pesada, o defraudar al lector, aunque el tema presentado sea interesante en sí mismo. El defecto más notable de una presentación que de otro modo sería aceptable es el *deus ex machina*. Antes de nuevos comentarios deseo dar un ejemplo concreto. Miremos la prueba del siguiente teorema, no completamente elemental **.

---

* Las secciones 4, 5 y 6 reproducen con ligeros cambios parte de mi ensayo «¿Con o sin motivación?», *American Mathematical Monthly*, v. 56, 1949, págs. 684-691.

** Se me disculpará que elija un ejemplo de mi propia obra. Ver G. Polya, «Proof of an inequality», *Proceedings of the London Mathematical Society* (2), v. 24, 1925, página lvii. El teorema probado se debe a T. Carleman.

*Si los términos de la secuencia* $a_1, a_2, a_3, \ldots$ *son números reales no negativos, no todos iguales a 0, entonces*

$$\sum_1^\infty (a_1\, a_2\, a_3 \ \ldots\ a_n)^{1/n} < e \sum_1^\infty a_n.$$

*Prueba.* Definir los números $c_1, c_2, c_3, \ldots$ en

$$c_1\, c_2\, c_3 \ \ldots\ c_n = (n+1)^n$$

para $n = 1, 2, 3, \ldots$ Usamos esta definición; luego, la desigualdad entre las medias aritméticas y geométricas (sección 8.6), y, finalmente, el hecho de que la secuencia que define $e$, cuyo término general es $[(k+1)/k]^k$, está aumentando. Obtenemos

$$\sum_1^\infty (a_1 a_2 \ldots a_n)^{1/n} = \sum_1^\infty \frac{(a_1 c_1\, a_2 c_2 \ldots a_n c_n)^{1/n}}{n+1}$$

$$\leqq \sum_1^\infty \frac{a_1 c_1 + a_2 c_2 + \ldots + a_n c_n}{n(n+1)}$$

$$= \sum_{k=1}^\infty a_k c_k \sum_{n=k}^\infty \frac{1}{n(n+1)}$$

(*d*)

$$= \sum_{k=1}^\infty a_k c_k \sum_{n=k}^\infty \left(\frac{1}{n} - \frac{1}{n+1}\right)$$

$$= \sum_{k=1}^\infty a_k \frac{(k+1)^k}{k^{k-1}} \frac{1}{k}$$

$$< e \sum_{k=1}^\infty a_k.$$

## 5. Justificación heurística

El punto de la derivación (*d*) es la definición de la secuencia $c_1, c_2, c_3 \ldots$ Este punto aparece desde el principio sin ninguna preparación, como un típico *deus ex machina*. ¿Qué le objetamos?

«Aparece como un conejo sacado de un sombrero.»

«Surge de la nada. Parece arbitrario. No tiene motivo o propósito visible.»

«Detesto andar en la oscuridad. Odio dar un paso cuando no veo la razón de que éste me lleve a la meta.»

«Quizá el autor conoce el propósito de este paso, pero yo no, y, por tanto, no puedo seguirle con confianza.»

«¡Mire, yo no estoy aquí sólo para admirarle! Deseo aprender cómo hacer los problemas yo solo. Pero no puedo ver cómo ha sido humanamente posible encontrar su... definición. «¿Qué puedo aprender en ella?», ¿Cómo podría encontrar tal... definición por mí mismo?»

«Este caso no es trivial. Parece crucial. Si pudiera ver que tiene algunas posibilidades de éxito o alguna justificación plausible provisional, entonces podría imaginar también cómo fue ideado y, en cualquier caso, podría seguir el razonamiento subsiguiente con más confianza y comprensión.»

Las primeras respuestas no son muy explícitas, las últimas son mejores y la última es la mejor. Revela que un lector o auditor inteligente desea dos cosas:

Primero, que el presente paso del argumento es correcto.

Segundo, ver que el presente paso es apropiado.

Un paso de un argumento matemático es apropiado si está esencialmente relacionado con el propósito, si nos acerca a la meta. Pero no es bastante que un paso sea apropiado: debe parecérselo al lector: Si el paso es sencillo, un paso trivial, de rutina, el lector puede fácilmente imaginar cómo se relacionará con el objetivo del argumento. Si el orden de presentación está planeado muy cuidadosamente, el contexto puede sugerir la relación del paso con el objetivo. Si, no obstante, el paso es visiblemente importante, pero su conexión con el objetivo no se ve en absoluto, aparece como un *deus ex machina* y el lector o auditor inteligente se siente comprensiblemente defraudado.

En nuestro ejemplo la definición de $c_n$ aparece como un *deus ex machina*. Sin embargo, este paso es, ciertamente, apropiado. En efecto, el argumento basado en esta definición prueba el teorema propuesto, y lo prueba con toda rapidez y claridad. El problema es que el paso en cuestión, aunque justificado al final, no lo aparece desde el principio.

Pero, ¿cómo podría el autor justificarlo desde el comienzo? La justificación completa lleva algún tiempo; es suministrada por la prueba total. Lo que necesita no es una justificación completa, sino una justificación *incompleta*, una *base provisional plausible*, sólo una indicación de que el paso tiene algunas probabilidades de éxito; en resumen, alguna justificación *heurística*.

## 6. La historia de otro descubrimiento

Es casi innecesario recordar al lector que las mejores historias no son verdaderas. Deben contener, no obstante, algunos elementos esenciales de verdad, si no no serían buenas. La siguiente es una presentación «algo racionalizada» de los pasos que me llevaron a la prueba dada en la sección 4. Esto es, la justificación heurística de los pasos sucesivos ha sido convenientemente acentuada.

El teorema probado en la sección 4 es sorprendente en sí mismo. Estaríamos menos sorprendidos si conociésemos cómo fue descubierto. Somos llevados a éste, naturalmente, tratando de probar lo siguiente: *Si la serie con números positivos*

$$a_1 + a_2 + a_3 + \ldots + a_n + \ldots$$

*es convergente, la serie*

$$a_1 + (a_1a_2)^{1/2} + (a_1a_2a_3)^{1/3} + \ldots + (a_1a_2a_3 \ldots a_n)^{1/n} + \ldots$$

*es también convergente.* Trataré de hacer hincapié en algunos motivos que pueden ayudarnos a encontrar la prueba.

(1) *Un teorema adecuado conocido.* Es natural empezar con las preguntas usuales.

*¿Cuál es la hipótesis?* Damos por supuesto que la serie $\Sigma a_n$ converge —que sus sumas parciales permanecen limitadas— que

$$a_1 + a_2 + \ldots + a_n \text{ no es grande.}$$

*¿Cuál es la conclusión?* Deseamos probar que la serie $\Sigma(a_1, a_2 \ldots a_n)^{1/n}$ converge que

$$(a_1a_2 \ldots a_n)^{1/n} \text{ es pequeño.}$$

*¿Conoce usted un teorema que pueda ser útil?* Lo que necesitamos es una relación entre la suma de $n$ cantidades positivas y su media geométrica.

*¿Ha visto usted algo de esta clase antes?* Si usted nunca ha oído hablar de la desigualdad entre las medias aritméticas y geométricas tiene una buena oportunidad de ello en esta ocasión:

$$(ag) \qquad (a_1a_2 \ldots a_n)^{1/n} \leqq \frac{a_1 + a_2 + \ldots + a_n}{n}.$$

Esta desigualdad muestra que $(a_1 a_2 \ldots a_n)^{1/n}$ es pequeña cuando $a_1 + a_2 + \ldots + a_n$ no es grande. Esto tiene tantos contactos con nuestro problema que apenas podemos resistir la tentación de aplicarlo:

$$\sum_{n=1}^{\infty} (a_1 a_2 \ldots a_n)^{1/n} \leqq \sum_{n=1}^{\infty} \frac{a_1 + a_2 + \ldots + a_n}{n}$$

(*a*)

$$= \sum_{k=1}^{\infty} a_k \sum_{n=k}^{\infty} \frac{1}{n}$$

—¡completo fracaso!—. Esta serie $\Sigma 1/n$ es divergente, la última línea de (*a*) carece de significado.

(2) *Aprendiendo del fracaso.* Es difícil admitir que nuestro plan estaba equivocado. Nos gustaría creer que al menos una parte de éste era correcto. Las preguntas útiles son: *¿Qué estaba equivocado en nuestro plan? ¿Qué parte de éste podremos salvar?*

La serie $a_1 + a_2 + \ldots + a_n + \ldots$ converge. Por tanto, $a_n$ es pequeña cuando $n$ es grande. Sin embargo, los dos lados de la desigualdad (*ag*) son diferentes cuando $a_1, a_2 \ldots a_n$ no son todas iguales, y pueden ser muy diferentes cuando $a_1 a_2 \ldots a_n$ son muy desiguales. En nuestro caso, $a_1$ es mucho más grande que $a_n$, y, así, puede haber un considerable abismo entre los dos lados de (*ag*). Esta es, probablemente, la razón de que nuestra aplicación de (*ag*) sea insuficiente.

(3) *Modificando el acercamiento.* El error estuvo en aplicar la desigualdad (*ag*) a las cantidades

$$a_1, a_2, a_3, \ldots a_n,$$

que son demasiado desiguales. ¿Por qué no aplicarlas a otras cantidades relacionadas que tengan más oportunidad de ser iguales? Podríamos intentar

$$1a_1, 2a_2, 3a_3, \ldots na_n.$$

¡Esta puede ser una idea! Podemos introducir tales factores compensatorios que van aumentando como 1, 2, 3, ... $n$. No obstante, no nos someteríamos más de lo necesario, reservándonos suficiente libertad de acción. Quizá podríamos considerar más generalmente las cantidades

$$1^{\lambda}a_1, 2^{\lambda}a_2, 3^{\lambda}a_3, \ldots n^{\lambda}a_n.$$

Por el momento dejaríamos $\lambda$ *indeterminada*, y al final elegimos el valor

más ventajoso. Este plan tiene tantos aspectos buenos que parece apropiado para la acción:

$$\sum_1^\infty (a_1 a_2 \ldots a_n)^{1/n} = \sum_1^\infty \frac{(a_1 1^\lambda \cdot a_2 2^\lambda \ldots a_n n^\lambda)^{1/n}}{(1 \cdot 2 \ldots n)^{\lambda/n}}$$

$$(b) \qquad = \sum_{n=1}^\infty \frac{a_1 1^\lambda + a_2 2^\lambda + \ldots + a_n n^\lambda}{n(n!)^{\lambda/n}}$$

$$= \sum_{k=1}^\infty a_k k^\lambda \sum_{n=k}^\infty \frac{1}{n(n!)^{\lambda/n}}.$$

Corramos con las dificultades. No podemos evaluar la última suma. Aun cuando superemos algunas de las dificultades, nos vemos obligados a trabajar con «ecuaciones crudas» (notación $\approx$, en lugar de $=$):

$$(n!)^{1/n} \approx ne^{-1},$$

$$\sum_{n=k}^\infty \frac{1}{n(n!)^{\lambda/n}} \approx e^\lambda \sum_{n=k}^\infty n^{-1-\lambda}$$

$$\approx e^\lambda \int_k^\infty x^{-1-\lambda}\, dx$$

$$= e^\lambda \lambda^{-1} k^{-\lambda}.$$

Al introducir esto en la última línea de ($b$) estamos muy cerca de probar

$$(b') \qquad \sum_1^\infty (a_1 a_2 \ \ldots \ a_n)^{1/n} \leqq C \sum_1^\infty a_k$$

donde $C$ es una constante, quizá $e^\lambda \lambda^{-1}$. Tal desigualdad probaría, desde luego, el teorema entre manos.

(4) *Mirando atrás* en el anterior razonamiento, nos vemos conducidos a repetir la pregunta: «¿Qué valor de $\lambda$ es el más ventajoso?». Probablemente, el $\lambda$, que hace de $e^\lambda \lambda^{-1}$ un mínimo. Podemos encontrar este valor por cálculo diferencial:

$$\lambda = 1.$$

Esto sugiere fuertemente que la elección más obvia es la más ventajosa: el factor compensador que multiplica $a_n$ debe ser $n^1 = n$, o alguna canti-

dad no muy diferente de $n$ cuando $n$ es grande. Esto puede conducir al valor sencillo de $C = e$ en $(b')$.

(5) *Más flexibilidad.* En nuestro anterior razonamiento $(b)$ dejamos a $\lambda$ indeterminado. Esto dio a nuestro plan cierta *flexibilidad:* el valor de $\lambda$ quedó a nuestra disposición. ¿Por qué no conceder a nuestro plan más flexibilidad? Podríamos haber dejado el valor compensador que multiplica $a_n$ bastante indeterminado; lo llamamos $c_n$, y podemos disponer de su valor final, cuando veamos con más claridad lo que necesitamos. Aventurémonos en esta nueva modificación de nuestro acercamiento original:

$$\sum_{1}^{\infty} (a_1 a_2 \dots a_n)^{1/n} = \sum_{n=1}^{\infty} \frac{(a_1 c_1 \cdot a_2 c_2 \dots a_n c_n)^{1/n}}{(c_1 c_2 \dots c_n)^{1/n}}$$

$$(c) \qquad \leqq \sum_{n=1}^{\infty} \frac{a_1 c_1 + a_2 c_2 + \dots + a_n c_n}{n(c_1 c_2 \dots c_n)^{1/n}}$$

$$= \sum_{k=1}^{\infty} a_k c_k \sum_{n=k}^{\infty} \frac{1}{n(c_1 c_2 \dots c_n)^{1/n}}.$$

¿Cómo elegir $c_n$? Esta es la pregunta crucial, cuya contestación no podemos posponer por más tiempo.

Primero: vemos con facilidad que el factor de proporcionalidad debe permanecer arbitrario. En efecto, la secuencia $cc_1, cc_2, \dots, cc_n, \dots$ conduce a las mismas secuencias que $c_1, c_2, \dots, c_n, \dots$

Segundo: nuestro anterior esfuerzo sugiere que tanto $c_n$ como $(c_1, c_2 \dots c_n)^{1/n}$ serán asintóticamente proporcionales a $n$:

$$c_n \sim Kn, (c_1 c_2 \dots c_n)^{1/n} \sim e^{-1} Kn = K'n.$$

Tercero: es muy conveniente que sepamos efectuar la suma

$$\sum_{n=k}^{n=\infty} \frac{1}{n(c_1 c_2 \dots c_n)^{1/n}}.$$

En este punto necesitamos todo el conocimiento anterior que tengamos sobre series. Si estamos familiarizados con la serie

$$\sum \frac{1}{n(n+1)} = \sum \left( \frac{1}{n} - \frac{1}{n+1} \right)$$

es una buena oportunidad que se nos ofrezca en esta coyuntura. Esta serie tiene la propiedad de que su suma es una sencilla expresión no sólo de $n = 1$ a $n = \infty$, sino también de $n = k$ a $n = \infty$ —¡una gran ventaja!—. Esta serie sugiere que la elección

$$(c_1 c_2 \dots c_n)^{1/n} = n + 1.$$

Ahora, visiblemente, $n + 1 \sim n$ para $n$ grande —¡un buen signo!—. ¿Qué hay sobre $c_n$? Como

$$c_1 c_2 \dots c_{n-1} c_n = (n+1)^n, \qquad c_1 c_2 \dots c_{n-1} = n^{n-1},$$

$$c_n = \frac{(n+1)^n}{n^{n-1}} = \left(1 + \frac{1}{n}\right)^n n \sim en;$$

la proporcionalidad asintónica con $n$ es una buena señal. Y el número $e$ aparece —¡una señal magnífica!

Elegimos $c_n$ y, tras esta elección, tomamos de nuevo la derivación (*d*) en la sección 4 con más confianza que antes.

Ahora podemos comprender cómo fue posible descubrir la definición de $c_n$ que aparece en la sección 4 como un *deus ex machina*. La derivación (*d*) se hizo también más comprensible. Aparece ahora como la única lograda en una cadena de ensayos consecutivos (*a*), (*b*), (*c*) y (*d*). Y el origen del teorema mismo está elucidado. Vemos cómo fue posible descubrir el papel del número *e*, que parecía tan sorprendente al principio.

## 7. Algunas indicaciones típicas

Acabamos de ver dos ejemplos. Primero, examinamos un «problema a encontrar» (en la sección 2), luego un «problema a probar» (en la sección 6)*. Una mayor variedad de ejemplos fueron necesarios para ilustrar con propiedad el papel de razonamiento plausible al idear un proyecto de la solución. En cualquier caso, por nuestros ejemplos pudimos desenmarañar unas cuantas circunstancias típicas indicadoras de la validez de un proyecto. Al tratar con otras circunstancias de esta clase apelaremos a cualquier experiencia que el lector tenga en la resolución de problemas matemáticos.

Al enumerar estas circunstancias indicadoras no tratamos de ser

---

* Para esta terminología ver *How to Solve It*, págs. 141-144.

exhaustivos. En algunos casos nos parecerá necesario distinguir entre problemas a encontrar y problemas a probar. En tales casos daremos dos formulaciones paralelas, y daremos primero la formulación relativa a los problemas a encontrar.

Consideremos una situación del razonamiento plausible que se presenta, naturalmente, al que está resolviendo un problema. Está usted ocupado con un interesante problema. Y tiene concebido un proyecto de solución, pero, en cierto sentido, no le gusta totalmente. Tiene sus dudas, no está completamente convencido de que su proyecto sea factible. Debatiendo este asunto consigo mismo, está usted, en efecto, examinando una conjetura:

*A. Este proyecto de solución dará resultado.*

A medida que examina usted su proyecto desde varios ángulos pueden ocurrírsele varios pros y contras. He aquí algunas indicaciones típicas, notables, que pueden hablar a favor de la conjetura *A*.

$B_1$. *Este proyecto tiene en cuenta todos los datos.*

Esta formulación se aplica a problemas a encontrar. Hay una formulación paralela que se aplica a los problemas a probar: *este proyecto tiene en cuenta todas las partes de la hipótesis.* Por ejemplo, la figura 16.3 combina todos los datos, y ésta es una buena señal. También la figura 16.2 contiene todos lo datos, pero hay una diferencia entre las dos figuras, que aclararemos en los próximos puntos.

$B_2$. *Este proyecto procura un enlace entre los datos y las incógnitas.*

Hay una formulación paralela que se refiere a problemas a probar: este proyecto proporciona una conexión entre la hipótesis y la conclusión. Por ejemplo, en la sección 6 (1) la desigualdad entre las medias aritméticas y geométricas ofreció crear una conexión entre la hipótesis y la conclusión. Y esta esperanza nos movió a trabajar con esa desigualdad. La figura 16.3 parece proporcionar un enlace más estrecho, y por eso parece más esperanzadora que la figura 16.2.

$B_3$. *Este proyecto tiene rasgos que son a menudo útiles en la resolución de problemas de esta clase.*

Por ejemplo, el proyecto comenzado en la figura 16.3 introduce en una etapa más madura (figura 16.4) la construcción de un triángulo. Esta es una buena señal, pues los problemas de construcción geométrica son a menudo reducidos a la construcción de triángulos.

$B_4$. *Este proyecto es semejante a otro que logró resolver un problema análogo.*

$B_5$. *Este proyecto logró resolver un caso particular del problema.*

Por ejemplo, usted tiene un proyecto para resolver un problema difícil que se refiere a una curva arbitrariamente cerrada. Realizar el

proyecto parece envolver mucho trabajo, y esto puede hacerle dudar. Pero usted observa que en el caso particular en que la curva cerrada es un círculo, el proyecto funciona muy bien y da un resultado correcto. Esto es una buena señal, y usted se siente animado.

$B_6$. *Este proyecto logró resolver una parte del problema* (al encontrar alguna de las incógnitas, o probando una conclusión más débil).

Esta lista no es de ninguna manera exhaustiva. Hay aún otras indicaciones y signos típicos que no necesitamos dar aquí. Además, sería inútil señalarlos sin ilustración adecuada *.

## 8. La inducción en la invención

La conjetura $A$ del estudiante, de que su proyecto de solución daría resultado, puede estar apoyada por una, dos, o más indicaciones, $B_1, B_2, B_3, \ldots$, de la clase señalada en la sección 7. Estas indicaciones se le pueden ocurrir al estudiante sucesivamente, intensificando cada una de ellas su confianza en su proyecto. Nuestras discusiones anteriores nos llevan a comparar tal proceso de resolución de un problema con el proceso inductivo en que un investigador, examinando una conjetura $A$, logra verificar consecuencias $B_1$, $B_2$, $B_3$, ... sucesivamente. Podemos también compararlo con el procedimiento legal en que el jurado que examina una acusación $A$ toma conocimiento de varias circunstancias corroboradoras $B_1$, $B_2$, $B_3$, ... sucesivamente. No debemos esperar ingenuamente la identidad de los tres procesos, sino que debemos examinar su semejanza o desemejanza con una mente abierta.

(1) Cuando el estudiante debate consigo mismo su proyecto para la resolución, éste es, generalmente, más fluido que rígido, más sentido que formulado. En efecto, sería una locura fijar el proyecto prematuramente. Un estudiante sensato no se limita a sí mismo a un proyecto rígido. Incluso en una etapa posterior, cuando el proyecto está más maduro, lo mantiene dispuesto para ser modificado, lo deja en una cierta flexibilidad, cuenta con dificultades imprevistas a las que podría verse obligado a adaptar su proyecto. Por tanto, cuando el estudiante examina la viabilidad de su proyecto examina un objeto cambiante y a veces transitorio.

Por otro lado, las conjeturas que el matemático o el naturalista investigan son, generalmente, más determinadas: están claramente formuladas, o al menos bastante próximas a una formulación clara.

---

* Cf. *How to Solve It*, págs. 212-221.

También el jurado tiene una conjetura bastante determinada que examinar: una acusación, cuyos términos han sido cuidadosamente establecidos en el proceso.

Observemos la sorprendente diferencia que separa la investigación de la viabilidad del proyecto del estudiante de la investigación inductiva de una conjetura física o matemática, o de la investigación de un delito: es la diferencia existente entre un objeto cambiable o transitorio y otro relativamente bien definido y determinado.

(2) Los procesos y actas de un tribunal de justicia son puestos en el expediente. La conjetura examinada por el naturalista y la evidencia recogida en favor o en contra de ella están también destinadas a un registro permanente. No ocurre lo mismo con la conjetura del estudiante referente a la viabilidad de su esquema, o con los signos que hablan en su favor o en su contra: su importancia es efímera. Son extremadamente importantes en tanto guían las decisiones del estudiante, pero cuando el trabajo entra en una nueva fase el proyecto mismo puede cambiar y entonces dichas indicaciones pierden casi todo su interés. Al final, cuando la solución es alcanzada y el problema está hecho, todos esos accesorios son abandonados. La forma final de la solución puede ser registrada, pero los proyectos cambiantes y los argumentos en favor o en contra de ellos son, en su mayoría, o totalmente, olvidados. Los edificios levantados permanecen a la vista, pero los andamios que fueron necesarios para levantarlos son retirados.

Anotemos este aspecto de la diferencia entre una investigación inductiva o judicial, por un lado, y la valoración de las perspectivas del trabajo del estudiante, por otro: la primera está destinada a un registro permanente, la segunda no.

(3) La conjetura $A$ y las indicaciones $B_1$, $B_2$, ... anotadas en la sección 7 pueden ser interpretadas con cierta amplitud. Después de las observaciones anteriores, en (1) y (2), no esperaremos que una interpretación agudamente definida sea aplicable con demasiada frecuencia. Hay alguna ventaja en empezar con tal interpretación. Consideremos la conjetura $A$ del estudiante y una indicación $B$ que la apoya:

*A.* Este proyecto para la solución dará resultado en su forma presente.

*B.* Este proyecto para la solución tiene en cuenta todos los datos.

Para describir la situación con más precisión añadimos: *Se sabe que todos los datos son necesarios.* Si esto es así,

$A$ implica $B$.

En efecto, si el proyecto funciona y da la solución correcta, debe usar todos los datos, cada uno de los cuales es necesario para la solución.

Ahora es importante visualizar claramente la situación: $A$ es una conjetura en que el estudiante está, naturalmente, interesado; $B$ es un enunciado que puede ser o no ser cierto. Examinemos ambas posibilidades.

(4) Casi todos los datos son necesarios para la solución, pero nuestro proyecto de solución no tiene en cuenta todos los datos; dicho proyecto, en su forma actual, no da resultado. (Podría hacerlo con una forma modificada.) Esto es, si $B$ es falsa, $A$ puede ser también falsa.

Ahora es importante observar que podríamos alcanzar esta conclusión por razonamiento formal. De hecho, hemos seguido aquí un patrón de razonamiento clásico y elemental (ya citado en la sección 12.1): el ***modus tollens*** del llamado silogismo hipotético:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ falsa}\end{array}}{A \text{ falsa.}}$$

(5) No obstante, si nuestro proyecto de solución toma en cuenta todos los datos, es natural considerar esta circunstancia como un buen signo, como un augurio favorable, como una predicción de que nuestro proyecto dará resultado. (Me imagino el descanso del estudiante cuando note que un dato que parecía en principio ser desdeñado por su proyecto entra más tarde en él, utilizándolo de una manera clara.) En resumen, si $B$ es cierta, $A$ se hace más digna de crédito.

Ahora es importante observar que, en efecto, podríamos haber llegado a esta conclusión usando simplemente nuestro patrón fundamental inductivo:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ verdadera}\end{array}}{A \text{ más digna de crédito.}}$$

(6) Consideremos ahora otra situación. Esta es semejante a la situación explicada en (3) y discutida en (4) y (5), aunque en el fondo es diferente. De nuevo estamos ante la conjetura $A$ y una indicación $B$, que la apoya. No obstante, la situación es ahora distinta (menos agudamente definida). $A$ y $B$ tienen los significados:

$A$. Este proyecto para la solución dará resultado (quizá en forma modificada).

*B*. Este proyecto para la solución tiene en cuenta todos los datos.

Para caracterizar la situación más plenamente tenemos que añadir: *Sospechamos fuertemente, aunque no lo sabemos definitivamente, que todos los datos son necesarios.*

Como dijimos anteriormente, $A$ es una conjetura a la que el estudiante «apuesta», y $B$ es un enunciado que puede, o no, ser verdadero. Tenemos que examinar las dos posibilidades.

Si $B$ resulta falsa hay un argumento en contra de $A$, pero no es totalmente decisivo. Dado que $B$ es falsa, nuestro proyecto no tiene en cuenta todos los datos; sin embargo, podemos perseverar en éste (si tenemos alguna base sólida para ello, aunque no esté explícita). Puede haber alguna razón (aún no clara) para esperar que una modificación de nuestro proyecto tendría en cuenta, finalmente, todos los datos. También la duda de que todos los datos pueden no ser necesarios podría tener alguna influencia.

Si $B$ resulta verdadera podemos tomar esta circunstancia como un signo alentador. En efecto, aunque $A$ no implique a $B$, y por ello $B$ no sea absolutamente cierta con $A$, aún puede ser que

$B$ con $A$ sea digna de crédito,

$B$ sin $A$ sea menos digna de crédito.

En tal caso, la verificación de $B$ aún cuenta como una especie de evidencia circunstancial de $A$. (Cf. sección 13.13 (5).)

(7) En las subsecciones anteriores (3), (4), (5) y (6) examinamos la indicación anotada en $B_1$ en la sección 7. (La llamamos $B$.) El examen de las otras indicaciones anotadas en la sección 7 (en $B_2$, $B_3$, ... $B_6$) mostrarán un cuadro semejante.

Como hemos visto, $A$ puede implicar a $B_1$, pero incluso si no es así, y $B_1$ no está necesariamente asociada con $A$, puede haber fuertes posibilidades de que $A$ vaya acompañada de $B_1$. La relación de $A$ a $B_2$ (o $B_3$, o $B_4$, ...) es de la misma naturaleza. Si el proyecto es bueno debe crear alguna conexión entre los datos y la incógnita (o entre la hipótesis y la conclusión); cf. $B_2$. No es absolutamente necesario que la solución sea parecida a la solución de algún problema semejante resuelto anteriormente, aunque las oportunidades de que esto sea así son bastante fuertes; cf. $B_3$, $B_4$. Si el proyecto da resultado para el problema total, deberá dar resultado para cualquier caso particular o parte del problema; cf. $B_5$, $B_6$.

Por tanto, si dudamos de $A$, pero hemos logrado éxito al observar

$B_1$ o $B_2$ o $B_3$ ..., podemos considerar razonablemente nuestra observación como una clase de evidencia inductiva o circunstancial en favor de $A$.

(8) Si, tras mucho trabajo, el naturalista sólo logra verificar unas cuantas consecuencias no muy sorprendentes de su conjetura, puede sentirse inclinado a abandonarla. Si demasiado poca evidencia es sometida contra el acusado, el tribunal puede retirar el caso. Si, después de un largo y extenuante esfuerzo, sólo se le han presentado al estudiante unas pocas y débiles indicaciones en pro de su proyecto, puede sentirse movido a modificar su proyecto, o incluso a abandonarlo.

Por otro lado, la conjetura del naturalista, del tribunal o del estudiante puede ser muy fortalecida si se verifican varias consecuencias, varios trozos de evidencia contra el acusado o varias indicaciones observadas. Sin embargo, más importante que el número puede ser la variedad. Tienen más peso consecuencias muy diferentes entre sí, testigos obviamente independientes, indicaciones que vienen de diferentes lados. (Cf. secciones 12.2, 13.11, 15.9.)

(9) A despecho de tales semejanzas, hay una diferencia considerable. La tarea del naturalista es recoger toda la evidencia que pueda en favor o en contra de su conjetura. La tarea del tribunal es examinar toda la evidencia de importancia sometida. Pero no es la tarea del estudiante que trabaja en un problema recoger tanta evidencia como pueda en favor o en contra de su proyecto para la resolución del mismo, o debatir tal evidencia hasta el final: su tarea es resolver el problema, por cualquier medio, siguiendo este proyecto para la solución, otro cualquiera, o ninguno.

Incluso un proyecto defectuoso puede servir al propósito del estudiante. Para resolver su problema tiene que movilizar y organizar las partes relevantes de su pasada experiencia. Trabajando con un proyecto defectuoso, pero con verdadero esfuerzo, el estudiante puede obtener algún punto de importancia que de otra manera habría permanecido escondido en el fondo. Esto puede darle un nuevo punto de partida. Un mal proyecto es frecuentemente útil en la resolución de problemas; puede conducir a un proyecto mejor.

(10) Dos personas que se enfrentan ante la misma evidencia pueden honestamente diferir, aunque trabajen con los mismos patrones de razonamiento plausible. Su concepción puede ser diferente. Toda mi concepción, mis razones no explícitas o inarticuladas, pueden influir en mi valoración de la evidencia experimental o judicial. Pueden influir aún más en mi valoración de las indicaciones en favor o en contra de mi

proyecto para la solución, y esto no sin razón. Es razonable que, tratrabajando en la solución de un problema, dé más peso que a otras condiciones a mis dictados internos y menos peso a razones claramente formuladas; estoy trabajando para sacar datos de importancia escondidos en el fondo del problema.

Incluso me parece que uno de los principales valores de un estudiante maduro es que sea capaz de tratar sagazmente con indicaciones en favor o en contra de la viabilidad de su proyecto, igual que un naturalista bien adiestrado con la evidencia experimental, o un ducho abogado con la evidencia legal.

## 9. Algunas palabras al profesor

Las matemáticas tienen muchos aspectos. Para muchos estudiantes me temo que aparecen como un conjunto de rígidas reglas, algunas de las cuales se aprenden de memoria antes de los exámenes finales, y todas ellas se pueden olvidar después. Para algunos maestros las matemáticas son como un sistema de rigurosas pruebas que, no obstante, hay que abstenerse de presentar en clase, presentando, en cambio, alguna charla no convincente, de la que se está algo avergonzado. Para un matemático activamente interesado en la investigación las matemáticas pueden parecer algunas veces un juego de adivinanzas: usted tiene que intuir un teorema matemático antes de probarlo, intuir la idea de la prueba antes de realizarla en sus detalles.

Para un filósofo de mente abierta toda inteligente adquisición de conocimientos aparecerá, a veces, como un juego de adivinanzas, creo yo. En la ciencia, como en la vida diaria, cuando nos enfrentamos a una nueva situación, nuestra primera respuesta suele ser una intuición. Puede ser que la intuición no dé en el blanco, pero lo intentamos, y, de acuerdo al grado de éxito obtenido, la modificamos más o menos. Así, después de varios ensayos y modificaciones, empujados por la observación y conducidos por la analogía, podemos llegar a una intuición más satisfactoria. El hombre de la calle no encuentra sorprendente que el naturalista trabaje de este modo. El conocimiento del naturalista puede estar mejor ordenado con vistas a seleccionar las analogías apropiadas, sus observaciones pueden ser más intencionadas y más cuidadosas, puede dar nombres fantásticos a sus intuiciones y llamarlas «generalizaciones tentativas», pero el naturalista debe adaptar su mente a la nueva situación mediante la intuición, como lo hace el hombre común. Y el hombre de la calle no se sorprende de oir que el naturalista intuye

como él mismo. Puede sorprenderle algo más a este hombre común oir que el matemático es también intuitivo. El resultado del trabajo creador del matemático es el razonamiento demostrativo, una prueba, pero la prueba se descubre por razonamiento plausible, es decir, por intuición.

Si esto es así, y yo lo creo, habrá un lugar para la intuición en la enseñanza de las matemáticas. La educación debe prepararnos para la invención, o, al menos, para el gusto de ella. En todo caso, la educación no debe suprimir los gérmenes inventivos en el estudiante. Un estudiante, algo interesado en el problema discutido en clase, *espera* un cierto tipo de solución. Si el estudiante es inteligente, prevé la solución hasta cierto punto: el resultado puede aparecer de esta o de la otra manera y hay oportunidades de obtenerlo por tal y tal procedimiento. El profesor intentará darse cuenta de lo que los estudiantes esperan, intentará descubrirlo, apuntará hacia lo que razonablemente esperan. Si el estudiante es menos inteligente y, sobre todo, si está aburrido, es probable que dé intuiciones aturdidas e irresponsables. El profesor debe mostrar entonces que la intuición en el ámbito matemático debe ser razonable, responsable, respetable. Mi consejo a los profesores de matemáticas de todos los grados puede resumirse en esta exclamación: *¡Enseñemos intuyendo!*

No he dicho que desdeñemos las pruebas. Por el contrario, debemos enseñar ambas cosas: probar e intuir; ambas clases de razonamiento: demostrativo y plausible. Más valioso que cualquier hecho o truco matemático, teorema o técnica, es para el estudiante aprender dos cosas:

Primero, distinguir una demostración válida de un intento válido, una prueba de una intuición.

Segundo, distinguir una intuición más razonable de otra menos razonable.

Hay casos especiales en que es más importante enseñar a intuir que a probar. Veamos la enseñanza del cálculo para estudiantes de ingeniería. (Por mi parte, tengo una larga y variada experiencia en esta clase de enseñanza.) Los ingenieros necesitan matemáticas, algunos de ellos tienen un saludable interés en éstas, pero no están educados para comprender $\varepsilon$-pruebas, no tienen tiempo para $\varepsilon$-pruebas, no están interesados en $\varepsilon$-pruebas. Enseñarles las reglas del cálculo como un dogma impuesto desde arriba no es educativo. Pretender que su prueba es completa cuando, de hecho, no lo es, no es honesto. Confesemos serenamente que sus pruebas son incompletas, pero demos serios fundamentos plausibles para los resultados probados de modo incompleto,

por medio de ejemplos y de la analogía. Luego no necesitamos avergonzarnos de pruebas fingidas, y algunos estudiantes recordarán su enseñanza después de los exámenes. Sobre la base de una larga experiencia me atrevo a decir que los estudiantes inteligentes de ingeniería son normalmente más accesibles y más agradecidos cuando se les presentan razones plausibles que pruebas estrictas.

He dicho que es deseable enseñar intuyendo, pero no es fácil hacerlo. No hay un método a toda prueba para intuir, y, en consecuencia, no puede haber ningún método a toda prueba para enseñar intuyendo. Puede ser que haya dicho algunas tonterías en lo anterior; pero espero haber evitado lo más tonto de todo, que es el pretender que tengo un método infalible para enseñar intuyendo.

Aun así no es imposible enseñar intuyendo. Tengo la esperanza de que algunos de los ejemplos explicados largamente y algunos de los ejercicios propuestos en lo anterior sirvan de últiles sugestiones. Estas sugestiones tienen una buena oportunidad de caer en campo fértil con profesores que tienen experiencia en la solución de problemas.

Tomemos, para muestra, el ejemplo examinado en las secciones 4 y 6. Las dos presentaciones, de las secciones 4 y 6, son muy diferentes. La diferencia más obvia es que una es corta y la otra larga. La diferencia más esencial es que una da prueba y la otra plausibilidades. Una está delineada para comprobar las *conclusiones demostrativas* justificando los pasos sucesivos. La otra es ordenada para iluminar los *motivos heurísticos* de ciertos pasos. La presentación de lo demostrativo sigue el modo usualmente aceptado desde Euclides; la presentación heurística es extremadamente desusada en publicaciones. Sin embargo, un profesor alerta puede usar ambas maneras de exposición. En realidad, podría utilizar una tercera presentación, si fuese necesario, que fuese un término medio entre las dos, considerando el tiempo disponible, el interés de los estudiantes y el resto de las condiciones en que trabaja *.

Este libro está principalmente dirigido a los estudiantes, con el propósito de desarrollar su habilidad, y a los lectores curiosos que quieran aprender algo sobre el razonamiento plausible y sus relaciones no banales con las matemáticas. Los intereses de los profesores creo que tampoco han sido desdeñados, aunque están expuestos más indirecta que directamente. Algún día espero llenar este hueco. Mientras tanto, reitero mi esperanza de que este libro pueda ser útil a los profesores, al menos a aquellos que tengan auténtica experiencia en la resolución de problemas.

---

* Para una presentación intermedia entre la de la sección 4 y la de la sección 6 ver G. H. HARDY, J. E. LITTLEWOOD y G. POLYA, *Inequalities*, págs. 249-250.

El problema es la existencia de tan pocos profesores de matemáticas que tengan tal experiencia. Incluso la mejor escuela de educación no ha logrado producir ese tipo maravilloso profesor con un adiestramiento tal en los métodos de enseñanza que pueda hacer comprender a sus estudiantes las cosas que él mismo no comprende bien.

## EJEMPLOS Y COMENTARIOS

**1.** *Para el profesor: algunos tipos de problemas.* Este libro ha sido escrito pensando en varias categorías de lectores: aquellos que desean comprender intuyendo, los que desean aprender intuyendo y los que desean enseñar intuyendo. Al lector de la última categoría se apela pocas veces de modo directo, pero un profesor alerta podría aprender algo de los ejemplos presentados en este libro y de su forma de presentación. Podría ver, por ejemplo, que hay maneras muy diferentes de la usual de proponer un problema. Deseo indicar unos cuantos tipos de problemas, los de «intuición-y-prueba», los de «prueba-de-consecuencias», los de «intuición-errónea» y los problemas «teoría-en-pequeña-escala». Todos estos tipos podrían ser utilizados por un profesor alerta para desafiar a sus estudiantes más inteligentes y descansar de la monotonía y de la rutina de los problemas que llenan los libros de texto.

*Intuición-y-prueba.* Los hechos matemáticos se intuyen primero y después se prueban; casi cada pasaje de este libro se esfuerza por mostrar que tal es el procedimiento normal. Si el aprendizaje de la matemática tiene algo que hacer en el descubrimiento de esta ciencia, hay que dar al estudiante alguna oportunidad de hacer problemas en que primero intuya y luego pruebe algunos hechos matemáticos a un nivel apropiado. Los libros de texto al uso no ofrecen tal oportunidad, los ejemplos 1.2, 5.1, 5.2, 7.1-7.6 (y muchos otros) lo hacen.

*Prueba-de-consecuencias.* Filósofos y no filósofos difieren en casi todo lo que respecta a la inducción, pero no hay mucha duda en que el procedimiento inductivo más frecuente consiste en examinar un enunciado general probando sus consecuencias particulares. Este procedimiento inductivo es de uso diario en la investigación matemática, y podría ser de uso diario también en la sala de clase, auténtico beneficio para los estudiantes. Ver sección 12.2 y ejemplos 12.3-12.6. Cf. ejemplo 6.

*Intuición-errónea.* Debemos adquirir alguna experiencia en intuición. Por nuestro contacto personal con las cosas reales sabemos que hay intuiciones respetables, que hay otras que conducen a error y que

incluso las propias intuiciones totalmente respetables pueden llevar a error. Para adquirir tal experiencia resolver los ejemplos 11.1-11.12.

*Teoría-en-pequeña-escala.* En casi todas las páginas de este libro se ha discutido algún problema relativamente elemental de manera que la discusión arrojase alguna luz sobre un punto que pudiera estar en conexión con otros problemas no tan elementales. Hay una razón para preferir esta «investigación-en-pequeña-escala»: un problema menos elemental podría mostrar el punto en cuestión en una escala más impresionante, pero exigiría una explicación mucho mayor y más conocimientos preliminares. No es demasiado fácil «reducir la escala»: pueden ser difíciles de encontrar problemas elementales que muestren con bastante claridad los rasgos relevantes del razonamiento plausible o inventivo. También es posible, pero menos fácil, idear problemas elementales que ilustren la actividad del científico a construir una teoría; los ejemplos siguientes 2, 3 y 4 ofrecen estos «problemas de teoría-en-pequeña-escala»; los ejemplos 5 y 6 son algo semejantes.

**2.** Un cuadrilátero es cortado en cuatro triángulos por sus dos diagonales. Llamamos a dos de estos triángulos «opuestos» si tienen un vértice común y ningún lado común. Probar los siguientes enunciados:

(*a*) El producto de las áreas de dos triángulos opuestos es igual al producto de las áreas de los otros dos triángulos opuestos.

(*b*) El cuadrilátero es un trapezoide solamente cuando tiene dos triángulos opuestos con la misma área.

(*c*) El cuadrilátero es un paralelogramo si sus cuatro triángulos son iguales en área.

**3.** (*a*) Probar el teorema siguiente: Un punto situado dentro de un triángulo equilátero dista de los tres lados $x$, $y$ y $z$, respectivamente; la altura del triángulo es $h$. Luego, $x + y + z = h$.

(*b*) Establecer y probar exactamente el teorema análogo en geometría sólida, referente a las distancias de un punto interior a las cuatro caras de un tetraedro regular.

(*c*) Generalizar los dos teoremas de modo que puedan aplicarse a cualquier punto de un plano o del espacio, respectivamente (y no sólo a puntos situados en el interior de un triángulo o de un tetraedro). Dar los enunciados exactos y también las pruebas.

**4.** Considerar las cuatro proposiciones siguientes, (I)-(IV), que no son necesariamente verdaderas.

(I) Si un polígono inscrito en un círculo es equilátero es también equiangular.

(II) Si un polígono inscrito en un círculo es equiangular es también equilátero.

(III) Si un polígono circunscrito a un círculo es equilátero es también equiangular.

(IV) Si un polígono circunscrito a un círculo es equiangular es también equilátero.

(*a*) Establecer cuáles de las cuatro proposiciones son verdaderas y cuáles son falsas, dando una prueba en cada caso.

(*b*) Si en vez de polígonos generales considerásemos sólo cuadriláteros ¿cuáles de las cuatro proposiciones serían verdaderas y cuáles falsas?

(*c*) ¿Qué ocurre con los pentágonos?

(*d*) ¿Puede usted intuir, o quizá incluso probar, enunciados más comprensivos: Esto explicaría sus observaciones (*b*) y (*c*).

**5.** Supongamos que $\alpha$, $\beta$, y $\gamma$ denotan los ángulos de un triángulo. Mostrar que

(*a*) $\operatorname{sen}\alpha + \operatorname{sen}\beta + \operatorname{sen}\gamma = 4\cos\frac{\alpha}{2}\cos\frac{\beta}{2}\cos\frac{\gamma}{2}$,

(*b*) $\operatorname{sen}2\alpha + \operatorname{sen}2\beta + \operatorname{sen}2\gamma = 4\operatorname{sen}\alpha\operatorname{sen}\beta\operatorname{sen}\gamma$,

(*c*) $\operatorname{sen}4\alpha + \operatorname{sen}4\beta + \operatorname{sen}4\gamma = -4\operatorname{sen}2\alpha\operatorname{sen}2\beta\operatorname{sen}2\gamma$.

**6.** Considerar una pirámide recta truncada de base cuadrada. Llamamos «sección media» a la intersección de la pirámide truncada con un plano paralelo a las dos bases y a la misma distancia de ambas. Llamamos «rectángulo intermedio» al rectángulo que tiene un lado igual a un lado de la base mayor y el otro lado igual a un lado de la base menor.

Cuatro amigos suyos están de acuerdo en que el volumen de la pirámide truncada es igual a la altura multiplicada por cierta área, pero no están de acuerdo en lo que respecta a este área y hacen cuatro proposiciones diferentes:

(I) la sección media,

(II) la media de las bases menor y mayor,

(III) la media de la base mayor, de la menor y de la sección media,

(IV) la media de la base mayor, de la menor y del rectángulo intermedio.

Supongamos que $h$ es la altura de la pirámide truncada, $a$ el lado de su base mayor y $b$ el lado de su base menor. Expresar cada una de las cuatro propuestas con notación matemática, decidir si es correcta o errónea y probar la respuesta.

**7.** «*Qui nimium probat, nihil probat.*» Es decir, si usted prueba demasiado no prueba nada. No sé en qué sentido entendería este proverbio clásico su propio autor; por mi parte, deseo explicar un sentido de la frase que me fue extremadamente valioso cuando empezaba a trabajar en matemáticas. La frase me recuerda uno de los signos más útiles para juzgar de la valía de un proyecto de solución.

He aquí la situación: deseamos probar una proposición. Esta proposición consta de una conclusión y una hipótesis que tiene varias cláusulas, y usted sabe que cada una de estas cláusulas es necesaria para la conclusión, es decir, que ninguna de ellas puede ser descartada sin hacer inválida la conclusión de la proposición. Usted ha concebido un proyecto de la prueba y se halla pesando las oportunidades de dicho proyecto. Si éste no utiliza todas las cláusulas tenemos que modificar el proyecto o rechazarlo: si da resultado y prueba la conclusión, aunque abandone esta o aquella cláusula de la hipótesis, prueba demasiado, es decir, algo falso, y, en consecuencia, no prueba nada.

He dicho que nuestro proyecto debe hacer entrar todas las cláusulas. Quiero decir que no basta con una mención de labios para afuera: nuestro plan debe hacer resaltar el uso esencial de cada cláusula en la prueba. La estructura que trata de apoyar la conclusión no puede hacerla surgir a menos que cada cláusula de la hipótesis tenga un sólido apoyo.

Puede ser muy difícil ingeniar un proyecto que haga entrar todas las cláusulas de la hipótesis. Por tanto, si un proyecto engarza todas estas cláusulas, podemos sentirnos satisfechos: hay un signo excelente, una fuerte indicación de que el proyecto dará resultado.

Para observaciones referentes sobre la solución de «problemas a encontrar», todos cuyos datos son necesarios, ver la sección 8 (3), (4), (5).

Si preferimos una sentencia francesa al proverbio latino, he aquí una: «La mariée est trop belle» («la recién casada es demasiado bella».) Me parece que no es necesarioam pliar esto; después de todo lo anterior el lector puede poner los detalles por su cuenta.

**8.** *Proximidad y credibilidad.* ¿Hasta dónde alcanza la solución? ¿Cuánto queda por hacer? Tales preguntas gravitan pesadamente sobre la mente del estudiante que tiene que terminar su tarea a tiempo fijo, pero están presentes en la mente de todo especialista en la resolución de problemas.

(1) Nos sentimos capaces de dar una respuesta a tales preguntas, hasta cierto punto, no de un modo preciso, claro, pero bastante correctamente en conjunto.

Por ejemplo, volvamos la vista a las figuras 16.1-16.5, y al proceso

de solución que representan. El familiarizado con la solución de problemas se dará cuenta de que la figura 16.3 está mucho más cerca de la solución que la figura 16.2; y se dará cuenta de que la solución está al alcance al llegar a la solución 16.4.

Al juzgar la proximidad de la solución podemos descansar sobre sentimientos inexpresados o sobre signos más claros. Cualquier signo indicador de que nuestro proyecto de solución da resultado puede interpretarse también como signo de progreso hacia la solución, y nos ayuda a estimar la distancia a que nos encontramos de ella.

(2) Consideremos la solución de un «problema a probar». El objetivo es probar (o refutar) cierto teorema. El estudiante puede confiar, o desconfiar, del teorema en que trabaja. Pero, si es un estudiante adelantado, estará preparado para revisar sus creencias. Y, así, las preguntas: «¿Es el teorema digno de crédito?, ¿Cómo es digno de crédito?, estarán siempre presentes en su mente, y unas veces más, otras menos, en el subconsciente. Si algo nuevo le viene a la vista, tendrá dos preguntas siempre preparadas: «¿Lo hace más digno de crédito, o menos?», «¿Nos acerca a la solución o no?». Su atención puede ser absorbida, naturalmente, por el nuevo hecho que no encuentra tiempo para formular su pregunta en palabras. Incluso puede darse a sí mismo la contestación sin palabras. Aun cuando dice «Buen signo», o «Mal signo», no se tomará la molestia de expresar largamente lo que quiere significar. ¿Se trata de un signo de la *proximidad* de la solución o de la *credibilidad* del teorema? Sin embargo, si es un buen especialista en la resolución de problemas, conoce bastante bien la importante diferencia entre proximidad y credibilidad, y esta diferencia se notará en su trabajo, en su manejo del problema.

(3) Recordar un nombre una vez conocido, aunque ahora olvidado, es una tarea más sencilla que resolver un problema matemático, aunque algo semejante. Podemos observar a menudo a la gente que recuerda un nombre, y aprenderemos algunas cosas interesantes de tales observaciones.

En una conversación su amigo necesita decirle un nombre (el de una tienda, el de un conocido o quizá el de un autor). Duda por un momento y dice: «Sólo un minuto y te lo diré», o «Espera un poco, lo recordaré dentro de un rato», o «Qué estúpido soy, no puedo acordarme en este momento, aunque estoy seguro de que lo recordaré en seguida, quizá hoy mismo». Como es obvio, su amigo trata de calibrar la posibilidad del nombre, trata de medir una especie de «distancia psicológica». Supongo que sus predicciones serán, aproximadamente, correctas, que su evaluación de la «distancia psicológica» será más o menos certera.

En una comparación con la resolución de problemas es también interesante observar que una persona que no puede recordar un nombre completamente puede hacerlo parcialmente o, quizá mejor expresado, puede ser capaz de recordar algunos rasgos del nombre. Usted puede oir decir a su amigo: «El nombre no es Battenberg —no es el esposo de la reina, naturalmente—, pero es un nombre alemán, de tres sílabas, muy parecido a Battenberg.» Y (yo he observado tales casos) su amigo puede estar en lo cierto en esas particularidades, aunque el nombre correcto no se le ocurra hasta después de unos días.

De modo bastante semejante, un matemático, a pesar de no haber resuelto su problema, puede predecir ciertos rasgos de la solución, de modo seguro.

Estos aspectos de la resolución de problemas, que son los más interesantes para el futuro matemático o para el profesor, no son realmente accesibles a los métodos de la psicología experimental. Quizá el recuerdo de un nombre, proceso semejante en algunos aspectos a la resolución de un problema matemático, podría ser fácilmente traído al horizonte de los experimentos psicológicos.

**9.** *Cálculo numérico y razonamiento plausible.* Aunque los números son considerados a menudo como símbolos de la mayor certeza que se puede conseguir, los resultados del cálculo numérico no son ciertos de ningún modo; son sólo plausibles. El cálculo numérico depende del razonamiento plausible de diversos modos.

(1) Tenemos que hacer un largo cálculo numérico. El resultado final se alcanza por una secuencia de pasos. Tenemos una buena oportunidad de realizar correctamente cada paso, pero hay muchos, y cada uno de ellos es una posibilidad de error; en consecuencia, el resultado final puede estar equivocado. ¿Cómo podríamos precavernos del error?

Calculemos el número deseado dos veces por procedimientos tan diferentes como sea posible. Si los dos cálculos dan resultados diferentes es seguro que al menos uno de ellos está equivocado, aunque pueden estarlo los dos. Si los dos cálculos están de acuerdo no es seguro que el resultado obtenido por dos veces sea correcto, pero puede serlo, y la concordancia es una señal de su corrección. El peso de esta señal depende de la diferencia entre los dos procedimientos utilizados.

Por ejemplo, el peso será muy pequeño si, habiendo hecho el cálculo una vez, lo repetimos inmediatamente sin cambiar el método: con el primer cálculo todavía fresco en la mente y en los dedos es fácil tropezar por segunda vez en el mismo lugar en que tropezamos la primera. Repetir el cálculo después de un rato es algo mejor, encargar hacerlo a

otra persona es considerablemente mejor y hacer el segundo cálculo con un método muy diferente es mucho mejor.

En efecto, si dos procedimientos bastante diferentes llegan al mismo resultado tenemos sólo dos conjeturas obvias: el resultado puede ser correcto, o el acuerdo puede deberse al azar. Si la probabilidad de una concordancia por mero azar es muy pequeña, la segunda de las dos conjeturas rivales es inverosímil en la misma medida; nos sentimos inclinados a rechazarla y, por ello, estamos dispuestos a dar más confianza a la primera conjetura, esto es, podemos confiar más en la corrección del resultado.

Cuanto más difieren los procedimientos de los dos cálculos, más realista es la evaluación y más sencilla de la probabilidad de su concordancia: la probabilidad de que los dos cálculos lleguen por azar a las mismas $n$ figuras es de $10^{-n}$; cf. sección 14.9 (3), ejemplo 14.11, pero también ejemplos 14.32 y 12.

En un cálculo en gran escala es una buena práctica hacer una coincidencia casual aún más improbable por la introducción de *controles múltiples*. Dos cálculos son realizados por métodos tan diferentes como sea convenientemente posible, de modo que puedan concordar si son correctos, no sólo en el resultado final, sino también en varios resultados intermedios. A medida que aumenta la concordancia se hace progresivamente más difícil atribuirla a mero azar, aunque, por supuesto, el azar nunca puede ser totalmente excluido y, por tanto, el resultado nunca puede estar totalmente garantizado.

(2) En lo anterior hemos asumido tácitamente que los dos cálculos cuyos resultados hemos comparado se sabe que son estrictamente equivalentes en teoría. Pero en matemáticas aplicadas tenemos que trabajar a menudo con aproximaciones, y podemos comparar resultados numéricos, que no concuerdan completamente, aun cuando todas las operaciones aritméticas sean correctas; nuestra esperanza es que concuerden aproximadamente. Por lo demás, la teoría del método de aproximación con el que trabajamos puede ser muy imperfectamente conocida. En tales circunstancias, el horizonte del razonamiento plausible es aún más amplio y se hace más azaroso. Cf. ejemplo 11.23.

(3) Dos matemáticos, $A$ y $B$, investigaron el mismo conjunto de nueve problemas combinatorios. No necesitamos conocer el tema de tales problemas (trataban del hipercubo de cuatro dimensiones), pero es importante saber que estaban ordenados en orden de dificultad creciente. Los dos primeros problemas son triviales, el tercero es fácil, el cuarto menos fácil, después van aumentando en complejidad y el último es el más difícil.

Tanto *A* como *B* resolvieron los problemas, pero sus resultados no coinciden. Aquí están sus soluciones de los nueve problemas:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *A:* | 1 | 1 | 4 | 6 | 19 | 27 | 47 | 55 | 78 |
| *B:* | 1 | 1 | 4 | 6 | 19 | 27 | 50 | 56 | 74. |

Esto es, *A* y *B* están de acuerdo en los seis primeros problemas, que son más fáciles, pero difieren en los últimos tres problemas, que son los más difíciles. En efecto, los métodos seguidos por ambos fueron muy diferentes.

*A* emprendió la resolución de cada uno de los nueve problemas independientemente de la de los otros. Su método para cada problema es algo diferente, y a medida que fue enfrentándose a problemas más difíciles también su método se hizo más complejo.

*B* trabajó en los problemas con un método uniforme. Su labor consta de dos partes. La primera, que es más difícil, es una preparación común para la solución de los nueve problemas. La segunda parte, más rutinaria, aplica el resultado de la primera parte a cada problema individualmente, de acuerdo con una regla uniforme. Los problemas parecen más difíciles tratados con el método de *A* que con el método de *B*.

Me parece que la situación descrita nos da una base razonable para confiar más en la solución dada por *B* que en la de *A*.

Puesto que los dos métodos, que son muy distintos, están de acuerdo respecto de los seis primeros problemas, y éstos son más fáciles en cierto modo, hay una buena razón para creer que la solución de estos problemas es correcta. En la solución de los tres primeros problemas no hay duda.

Dado que el resultado de la primera parte del trabajo de *B* es verificado por sus consecuencias en tres casos de nueve, y se verifica presumiblemente en tres casos más (no está ni verificado ni refutado en los tres casos restantes), hay una buena base para confiar en este resultado.

Si, por otro lado, la primera parte del trabajo de *B* fuera correcta (como presumiblemente lo es) podría sólo errar en la segunda parte, más rutinaria, al tratar los últimos tres problemas. *A* tiene las mayores dificultades al enfrentarse con los problemas. Y, por eso, parece tener (*A*) mayores posibilidades de error que *B**.

El caso que acabamos de examinar es más que especial, pero muestra que hay nuevas posibilidades al explorar los patrones de razonamiento plausible. Por ejemplo, puede ser un trabajo compensador expresar el

* Ver G. POLYA, «Sur les types des propositions composées», *Journal of Symbolic Logic*, vol. 5 (1940), págs. 98-103.

argumento plausible, que acabamos de presentar, en fórmulas del cálculo de probabilidad, tan apropiadamente como sea posible.

**10.** Añadir una columna de diez números de seis cifras, que empiecen así, por ejemplo:

1 5 9 6 . 0 3
1 6 4 6 . 0 7
1 7 8 1 . 1 0
. . . . . .

Describir diversos procedimientos para hacerlo.

**11.** Se llama «paso elemental» a la adición de un número de un dígito que vemos escrito a un número de dos dígitos que tenemos en la mente; se incluye, no obstante, la posibilidad (que hace más sencillo el paso) de que el segundo número esté también escrito, o que tenga solamente un dígito. ¿Cuántos pasos elementales se necesitan para realizar la adición mencionada en el ejemplo 10, de la forma más usual?

**12.** Al hacer un cálculo obtenemos primero dos números de nueve cifras, y luego el resultado final de su diferencia, que resulta ser un número de tres cifras. Otro procedimiento de cálculo obtiene el mismo resultado de tres cifras, de la diferencia de dos números de siete cifras. Calcular la probabilidad de que una concordancia semejante sea debida al azar, usando la fórmula dada en el ejemplo 9 (1).

**13.** *Demostración formal y razonamiento plausible.* Tenemos que hacer un largo cálculo numérico. El resultado final se alcanza en una secuencia de siete pasos y debemos examinar si cada paso es correcto. Cada paso particular (como la suma $3 + 7$ o la multiplicación $3 \times 7$) es tan sencillo y familiar que no podemos equivocarnos en circunstancias un poco favorables, cuando prestamos una atención total. Aun así, como todo el mundo, podemos cometer errores en el cálculo. Después de haber realizado los pasos sucesivos con bastante cuidado, no debemos confiar en el resultado final sin haberlo comprobado.

Realizamos una larga demostración matemática, que se supone que hemos descompuesto en varios pasos para comprobar perfectamente cada uno de ellos, y la conclusión final será verdadera si cada paso es correcto. Pero podemos haber cometido un error. Después de haber comprobado cuidadosamente los pasos sucesivos ¿podemos confiar en la conclusión final? No más, y quizá menos, que el resultado final de un largo cálculo.

De hecho, un matemático que ha comprobado los detalles de una

demostración paso a paso y los ha encontrado bien a todos puede estar todavía satisfecho. Necesita para satisfacerse algo más que la corrección de los detalles. ¿El qué?

Necesita *comprender* la demostración. Después de haber examinado la prueba paso a paso le quedan aún dudas: revisa, vuelve a realizar las operaciones, formula de nuevo y vuelve a ordenar los pasos hasta lograr agrupar los detalles en un todo comprensible. Sólo cuando ha hecho todo esto empieza a confiar en la prueba.

No me atrevería a analizar lo que constituye la «comprensibilidad». Algunos dicen que está basada en la intuición y consideran a ésta como una percepción del todo, y una agrupación de los detalles dentro de un conjunto armónicamente ordenado. No me atrevería a contradecir esto, aunque tengo algunas dudas *. Sin embargo, deseo llamar la atención hacia un punto fuertemente sugerido por los efectos y exámenes de este libro.

Alguna práctica puede convencernos de que la analogía y los casos particulares pueden ser valiosos, tanto para encontrar como para comprender las demostraciones matemáticas. El proyecto general, o partes considerables de una prueba, pueden ser sugeridas o aclaradas por la analogía. Los casos particulares pueden sugerir una prueba (ver, por ejemplo, la sección 3.17); por otro lado, podemos comprobar una prueba ya formulada, observando cómo funciona en un caso particular crítico o familiar. Sin embargo, la analogía y los casos particulares son las fuentes más abundantes de argumentos plausibles: quizá no ayuden a formar y a hacer más comprensible el argumento demostrativo, pero aumentan nuestra *confianza* en él. Y, así, nos vemos llevados a sospechar que una buena parte de *nuestra confianza en los razonamientos demostrativos puede proceder del razonamiento plausible.*

---

* El significado de la intuición y su papel en la agrupación de detalles no ha sido bien explicado generalmente. Es digno de ser observado, no obstante, que Descartes, con quien empieza el uso moderno del término «intuición», explica ambos puntos, de modo bastante notable, en la tercera y la séptima de sus reglas para la dirección de la mente. Ver *Oeuvres*, editadas por Adam y Tannery, vol. 10, págs. 368-370 y 387-388.

# SOLUCIONES

# Soluciones

## CAPITULO PRIMERO

**1.** Los primos que terminan en 1.

**2.** [Stanford, 1948]

$$(n^2 + 1) + (n^2 + 2) + \ldots + (n + 1)^2 = n^3 + (n + 1)^3$$

Los términos del lado izquierdo están en progresión aritmética.

**3.** $1 + 3 + \ldots + (2n - 1) = n^2$.

**4.** 1, 9, 36, 100, ... son cuadrados. Véase *How to Solve It*, pág. 104.

**5.** [Stanford, 1949] $\left(\frac{n+1}{2}\right)^2$ o $\left(\frac{n+1}{2}\right)^2 - \frac{1}{4}$, según que $n$ sea impar o par. Una ley uniforme para ambos casos: el entero más próximo a $(n + 1)^2/4$.

**6.** Primera pregunta: Sí. Segunda pregunta: No; 33 no es primo.

**7.** No para usted, si tiene alguna experiencia con primos [ejemplos 1, 6, 9]. En efecto, (1) puede probarse (como caso particular de un teorema de Kaluza, *Mathematische Zeitschrift*, vol. 28, 1928, págs. 160-170), y (2) refutado: el próximo coeficiente (de $x^7$) es $-3447 = -3 \cdot 3 \cdot 383$. El «cálculo formal» tiene un claro significado. Haciendo

$$\left(\sum_0^\infty n!x^n\right)^{-1} = \sum_0^\infty u_n x^n,$$

definimos $u_0 = 1$ y $u_1, u_2, u_3, \ldots$ de modo recurrente por las ecuaciones

$$0!u_n + 1!u_{n-1} + 2u!_{n-2} + \ldots + (n - 1)!u_1 + n!u_0 = 0$$

para $n = 1, 2, 3, \ldots$

**8.** Sobre la base de los datos observados es bastante razonable sospechar que $A_n$ es positivo y aumenta con $n$. Pero esta conjetura está completamente equivocada. Por métodos más avanzados (cálculo integral, o teoría de funciones analíticas para una variable compleja) podemos probar que, para $n$ grande, el valor de $A_n$ es aproximadamente $(-1)^{n-1}\,(n-1)!(\log n)^{-2}$.

**10.** En el caso $2n = 60$ tenemos que hacer nueve o siete intentos ($p = 3, 5, 7, 11, 13, 17, 19, 23, 29$ o $p' = 31, 37, 41, 43, 47, 53, 59$), según que sigamos el primero o el segundo procedimiento. Es verosímil que para más altos valores de $n$ habrá aún mayor diferencia entre el número de intentos, en favor del segundo procedimiento.

Sin solución: **9, 11, 12, 13, 14.**

## CAPITULO II

**1.** Creo que $C$ o $D$ es la «generalización correcta» y $B$ «sobrepasa la marca». $B$ es demasiado general para una sugerencia específica. Usted puede elegir entre $C$ o $D$; la elección depende de su concepción. Sin embargo, $C$ y $D$ sugieren empezar con las ecuaciones lineales y conducen eventualmente al siguiente proyecto: expresar dos incógnitas en términos de la tercera incógnita a partir de las dos primeras ecuaciones (¡lineales!) y, sustituyendo estas expresiones en la primera ecuación, obtener una ecuación cuadrática para la tercera incógnita. (En el presente caso $A$, ¿puede usted expresar dos incógnitas cualesquiera a partir de las dos primeras ecuaciones?) Hay dos soluciones:

$$(x, y, z) = (1, -2, 2), \qquad (29/13, -2/13, 2).$$

**2.** Haciendo girar sobre su eje a la pirámide 180° coincide consigo misma. La generalización correcta de esta pirámide es un sólido que tiene un *eje de simetría* de esta clase y la solución *más sencilla* es un plano que pasa a través del eje y el punto dado. (Hay una infinidad de soluciones; por continuidad, podemos prescribir una línea recta a través de la cual pase el plano que bisecciona la pirámide.) Nótese que una pirámide regular con base pentagonal no admite una solución sencilla comparable. Confróntese *How to Solve It*, págs. 98-99.

**3.** $A$ es un caso especial de $B$ si admitimos en $B$ que $P$ puede coincidir con $O$; sin embargo, los dos problemas son equivalentes: los planos pedidos en $A$ y en $B$ son paralelos el uno al otro; de esta manera, la solución de cada problema implica la del otro.

El problema más general, $B$, es más accesible, dado que $P \neq 0$: elegir $Q$ y $R$, en las otras dos líneas, de modo que $OP = OQ = OR$. El plano que pasa por $P$, $Q$ y $R$ satisface las condiciones del problema. Por tanto, si se propone $A$, hay una ventaja al pasar al más general, $B$.

**4.** $A$ es un caso especial de $B$ (para $p = 1$); no obstante, los dos problemas son equivalentes: la sustitución $x = yp^{1/2}$ reduce $B$ a $A$.

El problema más general $B$ es más asequible: diferenciar la integral

$$\int_{-\infty}^{\infty} (p + x^2)^{-1}\, dx$$

dos veces respecto del parámetro $p$. Por tanto, si se propone $A$, tenemos una ventaja al pasar al más general, $B$.

Obsérvese una situación paralela en el ejemplo 3.

**6.** El caso extremo, especial, en que uno de los círculos degenera en un punto es más asequible y podemos reducirlo al caso general. En efecto, una tangente exterior común a dos círculos permanece paralela consigo misma cuando ambos radios aumentan en la misma cantidad, y una tangente común interna permanece paralela consigo misma cuando uno de los radios aumenta y el otro disminuye en la misma cantidad. En los dos casos podemos reducir uno de los círculos a un punto sin cambiar la dirección de la tangente común.

**8.** El caso especial en que uno de los lados del ángulo situado en la circunferencia pasa a través del centro del círculo es importante. A partir de dos de tales ángulos podemos combinar el ángulo general en la circunferencia, por adición o por sustracción. (Este es el *quid* de la prueba clásica; EUCLIDES, III, 20.) Ver en *How to Solve It*, págs. 166-170, un sorprendente ejemplo de un caso especial de esta clase.

**12.** Si dos líneas rectas en un plano son cortadas por tres líneas paralelas, los segmentos correspondientes son proporcionales. Esto ayuda a probar el teorema análogo (más difícil) de geometría sólida; véase EUCLIDES, XI, 17.

**13.** Las diagonales de un paralelogramo se cortan en su punto medio común.

**14.** La suma de dos lados cualesquiera de un triángulo es mayor que el tercer lado. El más sencillo de los dos teoremas análogos (EUCLIDES, I, 20) se utiliza en la prueba del más difícil (EUCLIDES, XI, 20).

**15.** El paralelepípedo, el paralelepípedo rectangular (caja), el cubo,

el plano que biseccione a un ángulo diedro. *Los planos que biseccionan los seis ángulos diedros de un tetraedro se encuentran en un punto que es el centro de la esfera inscrita en el tetraedro.*

**16.** El prisma, el prisma recto, la esfera. *El volumen de una esfera es igual al volumen de una pirámide, cuya base tiene la misma área que la superficie de la esfera y cuya altura es el radio.*

**17.** Llamamos a una pirámide *pirámide isósceles* si todas las aristas que empiezan en el vértice son iguales. Todas las caras laterales de una pirámide isósceles son triángulos isósceles. *La base de una pirámide de esta clase está inscrita en un círculo y su altura pasa por el centro de este círculo.* Cf. ejemplo 9.26.

**22.** Sí. Intercambiando $x$ y $-x$ no cambiamos $x^2$ o el producto que representa $(\operatorname{sen} x)/x$ de acuerdo con $E$.

**23.** Predicción: de $E$ se sigue

$$\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\left(1 - \frac{x^2}{16\pi^2}\right)\ldots = -\frac{\pi^2}{x(x+\pi)}\,\frac{\operatorname{sen} x - \operatorname{sen}\pi}{x-\pi}$$

y, por tanto, para $x \to \pi$, por la definición de la derivada,

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\left(1 - \frac{1}{16}\right)\ldots = -\frac{\pi^2}{\pi \cdot 2\pi}\cos\pi = \frac{1}{2}.$$

Verificación:

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right)\ldots\left(1 - \frac{1}{n^2}\right) = \frac{1\cdot 3}{2\cdot 2}\cdot\frac{2\cdot 4}{3\cdot 3}\cdot\frac{3\cdot 5}{4\cdot 4}\ldots$$

$$\frac{(n-1)(n+1)}{n\cdot n} = \frac{1}{2}\,\frac{n+1}{n} \to \frac{1}{2}.$$

**24.** 1/6. Como en el ejemplo 23, o el caso especial $k = 2$ del ejemplo 25.

**25.** Predicción: si $k$ es un entero positivo,

$$\prod_{n=1}^{k-1}\left(1 - \frac{k^2}{n^2}\right)\prod_{n=k+1}^{\infty}\left(1 - \frac{k^2}{n^2}\right) = \lim_{x\to k\pi}\frac{k^2\pi^2}{x(x+k\pi)}\,\frac{\operatorname{sen} x}{k\pi - x}$$

$$= (1/2)(-\cos k\pi) = (-1)^{k-1}\,{}^2.$$

Verificación: para $N \geqq k+1$,

$$\prod_{n=1}^{k-1} \frac{(n-k)(n+k)}{n\,n} \prod_{n=k+1}^{N} \frac{(n-k)(n+k)}{n\,n}$$

$$= \frac{(-1)^{k-1}(k-1)!(N-k)! \cdot (N+k)!/(k!2k)}{(N!/k)^2}$$

$$= \frac{(-1)^{k-1}}{2} \frac{(N-k)!(N+k)!}{N!\,N!}$$

$$= \frac{(-1)^{k-1}}{2} \frac{(N+1)(N+2)\ldots(N+k)}{N \quad (N-1)\ldots(N-k+1)} \to \frac{(-1)^{k-1}}{2}$$

a medida que $N$ tiende a $\infty$.

**26.** $\pi/4$, el área del círculo con diámetro 1. A partir de $E$, para $x = \pi/2$,

$$\frac{2}{\pi} = \left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{16}\right)\left(1 - \frac{1}{36}\right)\left(1 - \frac{1}{64}\right)\left(1 - \frac{1}{100}\right)\ldots$$

$$= \frac{1 \cdot 3}{2 \cdot 2} \cdot \frac{3 \cdot 5}{4 \cdot 4} \cdot \frac{5 \cdot 7}{6 \cdot 6} \cdot \frac{7 \cdot 9}{8 \cdot 8} \cdot \frac{9 \cdot 11}{10 \cdot 10} \ldots$$

Esta fórmula, debida a Wallis (1616-1703), fue bien conocida de Euler. Hay otra manera de exponer la fórmula de Wallis:

$$\frac{1}{\pi} = \lim_{n \to \infty} \left( \frac{1 \cdot 3 \cdot 5 \ldots (2n-1)}{2 \cdot 4 \cdot 6 \ldots \quad 2n} \right)^2 n.$$

**27.** $x = \pi z$ en el ejemplo 21 y la definición de un producto infinito.

**28.** Sí. A partir del ejemplo 27,

$$\frac{\operatorname{sen} \pi(z+1)}{\pi} = \lim_{n \to \infty} \frac{z+n+1}{z-n} \cdot \frac{(z+n)\ldots(z+1)\,z(z-1)\ldots(z-n)}{(-1)^n\,(n!)^2}$$

**29.** A partir del ejemplo 27 y del ejemplo 26,

$$\cos \pi z = \operatorname{sen} \pi(-z + 1/2)$$

$$= \pi \lim \frac{(-z+n+\frac{1}{2})\ldots(-z+\frac{3}{2})(-z+\frac{1}{2})(-z-\frac{1}{2})\ldots(-z+\frac{1}{2}-n)}{(-1)^n(n!)^2}$$

$$= \lim \frac{(2n-1-2z)\ldots(3-2z)(1-2z)(1+2z)(3+2z)\ldots(2n-1+2z)}{(2n-1)\ldots 3\cdot 1\cdot 1\cdot 3\ldots(2n-1)}$$

$$\cdot \lim \frac{-z+n+\frac{1}{2}}{n} \cdot \pi \lim \left(\frac{1\cdot 3\cdot 5\ldots(2n-1)}{2\cdot 4\cdot 6\ldots 2n}\right)^2 n$$

$$= \left(1 - \frac{4z^2}{1}\right)\left(1 - \frac{4z^2}{9}\right)\left(1 - \frac{4z^2}{25}\right)\ldots$$

**30.** Sí. A partir de $E$ y del ejemplo 29,

$$\frac{2 \operatorname{sen} \pi z/2}{\pi z} \cdot \cos \pi z/2 = \left(1 - \frac{z^2}{4}\right)\left(1 - \frac{z^2}{16}\right)\left(1 - \frac{z^2}{36}\right)\ldots$$

$$\cdot \left(1 - \frac{z^2}{1}\right)\left(1 - \frac{z^2}{9}\right)\left(1 - \frac{z^2}{25}\right)\ldots$$

$$= \left(1 - \frac{z^2}{1}\right)\left(1 - \frac{z^2}{4}\right)\left(1 - \frac{z^2}{9}\right)\left(1 - \frac{z^2}{16}\right)\ldots$$

$$= \frac{\operatorname{sen} \pi z}{\pi z}.$$

**31.** Predicción: para $x = \pi$ el ejemplo 29 da por resultado cos $\pi = -1$.

Verificación: el producto de los primeros $n$ factores

$$\frac{-1\cdot 3}{1\cdot 1}\cdot\frac{1\cdot 5}{3\cdot 3}\cdot\frac{3\cdot 7}{5\cdot 5}\cdot\frac{5\cdot 9}{7\cdot 7}\cdots\frac{(2n-3)(2n+1)}{(2n-1)(2n-1)} = -\frac{2n+1}{2n-1} \to -1.$$

**32.** Predicción: el ejemplo 29 da como resultado cos $2\pi = 1$.
Verificación: como en el ejemplo 31, o los ejemplo 31 y 35.

**33.** Predicción: para $x = n\pi (n = 1, 2, 3, \dots)$ del ejemplo 29 resulta

$$\left(1 - \frac{4n^2}{1}\right)\left(1 - \frac{4n^2}{9}\right)\left(1 - \frac{4n^2}{25}\right) \dots = \cos \pi n = (-1)^n.$$

Verificación: Partiendo del $\cos 0 = 1$ y del ejemplo 35, o directamente como en el ejemplo 31.

**34.** Sí. Como en el ejemplo 22.

**35.** Sí. Por el resultado, o el método, del ejemplo 28.

**36.** $$1 - \operatorname{sen} x = 1 - \cos\left(\frac{\pi}{2} - x\right) = 2 \operatorname{sen}^2\left(\frac{\pi}{4} - \frac{x}{2}\right)$$

$$= \left(\frac{\operatorname{sen} \pi(1 - 2z)/4}{\operatorname{sen} \pi/4}\right)^2;$$

hacemos $x = \pi z$. Por el ejemplo 27,

$$\frac{\operatorname{sen} \pi(1 - 2z)/4}{\operatorname{sen} \pi/4}$$

$$= \lim \frac{n + (1-2z)/4}{n + 1/4} \dots \frac{1 + (1-2z)/4}{1 + 1/4} \frac{(1 - 2z)/4}{1/4} \frac{-1 + (1-2z)/4}{-1 + 1/4}$$

$$\dots \frac{-n + (1 - 2z)/4}{-n + 1/4}$$

$$= \lim \frac{4n + 1 - 2z}{4n + 1} \dots \frac{5 - 2z}{5} \frac{1 - 2z}{1} \frac{3 + 2z}{3} \dots \frac{4n - 1 + 2z}{4n - 1}$$

$$= \left(1 - \frac{2z}{1}\right)\left(1 + \frac{2z}{3}\right)\left(1 - \frac{2z}{5}\right) \dots \left(1 + \frac{2z}{4n - 1}\right)\left(1 - \frac{2z}{4n + 1}\right) \dots$$

**37.** Pasando a los logaritmos y diferenciando en el ejemplo 21 o el 27. El significado preciso del lado derecho es

$$\lim_{n \to \infty} \left(\frac{1}{x + n\pi} + \dots + \frac{1}{x + \pi} + \frac{1}{x} + \frac{1}{x - \pi} + \dots + \frac{1}{x - n\pi}\right).$$

**38.** Por el ejemplo 37

$$\cot x = \frac{1}{x} + \sum_{n=1}^{\infty}\left(\frac{1}{x+n\pi} + \frac{1}{x-n\pi}\right)$$

$$= \frac{1}{x} + \sum_{n=1}^{\infty}\frac{2x}{x^2 - n^2\pi^2}$$

$$= \frac{1}{x} - 2x\sum_{n=1}^{\infty}\left(\frac{1}{n^2\pi^2} + \frac{x^2}{n^4\pi^4} + \frac{x^4}{n^6\pi^6} + \ldots\right).$$

Establecemos

$$y = \cot x = \frac{1}{x} + a_1x + a_2x^3 + a_3x^5 + \ldots$$

Luego, expresando el coeficiente de $x^{2n-1}$, encontramos

$$S_{2n} = 1 + \frac{1}{2^{2n}} + \frac{1}{3^{2n}} + \frac{1}{4^{2n}} + \frac{1}{5^{2n}} + \ldots = -\frac{a_n\pi^{2n}}{2}$$

para $n = 1, 2, 3, \ldots$ Para encontrar los coeficientes $a_1, a_2, a_3, \ldots$ utilizamos la ecuación diferencial

$$y' + y^2 = -1.$$

Sustituyendo $y$ e $y^2$ por sus desarrollos y comparando los coeficiente, de los valores semejantes de $x$, obtenemos relaciones entre los coeficientes $a_1, a_2, a_3, \ldots$, que podemos examinar más convenientemente en este cuadro:

| | $x^{-2}$ | 1 | $x^2$ | $x^4$ | $x^6$ | ... |
|---|---|---|---|---|---|---|
| $y'$ | $-1$ | $a_1$ | $3a_2$ | $5a_3$ | $7a_4$ | ... |
| $y^2$ | 1 | $2a_1$ | $2a_2$ | $2a_3$ | $2a_4$ | ... |
| | | | $a_1^2$ | $2a_1a_2$ | $2a_1a_3$ | ... |
| | | | | | $a_2^2$ | ... |
| | | | | | | ... |
| | 0 | $-1$ | 0 | 0 | 0 | ... |

Cf. el ejemplo 5.1. Obtenemos así las relaciones

$$3a_1 = -1, \quad 5a_2 + a_1^2 = 0, \quad 7a_3 + 2a_1a_2 = 0, \ldots$$

y de aquí, sucesivamente,

$$S_{2n} = -\frac{a_n \pi^{2n}}{2} = \frac{\pi^2}{6}, \frac{\pi^4}{90}, \frac{\pi^6}{945}, \frac{\pi^8}{9450}, \ldots$$

para

$$n = 1, 2, 3, 4, \ldots$$

**39.** El método del ejemplo 37 y del ejemplo 38 aplicado al resultado del ejemplo 36. Hacemos ahora

$$y = \cot\left(\frac{\pi}{4} - \frac{x}{2}\right) = b_1 + b_2 x + b_3 x^2 + b_4 x^3 + \ldots .$$

Luego,

$$T_n = 1 + \frac{(-1)^n}{3^n} + \frac{1}{5^n} + \frac{(-1)^n}{7^n} + \frac{1}{9^n} + \frac{(-1)^n}{11^n} + \ldots = \frac{b_n \pi^n}{2^{n+1}}.$$

Ahora $y$ satisface a la ecuación diferencial

$$2y' = 1 + y^2,$$

o cual (obsérvese que $b_1 = 1$) da el siguiente cuadro:

| | $1$ | $x$ | $x^2$ | $x^3$ | $x^4$ | ... |
|---|---|---|---|---|---|---|
| | $1$ | | | | | |
| $y^2$ | $1$ | $2b_2$ | $2b_3$ | $2b_4$ | $2b_5$ | ... |
| | | | $b_2^2$ | $2b_2 b_3$ | $2b_2 b_4$ | ... |
| | | | | | $b_3^2$ | ... |
| | | | | | | ... |
| $2y'$ | $2b_2$ | $4b_3$ | $6b_4$ | $8b_5$ | $10b_6$ | ... |

De aquí obtenemos, primero, las relaciones $2b_2 = 2$, $4b_3 = 2b_2$, $6b_4 = 2b_3 + b_2^2$, ..., y luego los valores para

$$n = 1, 2, 3, 4, 5, 6, \ldots$$

$$T_n = \frac{\pi}{4}, \frac{\pi^2}{8}, \frac{\pi^3}{32}, \frac{\pi^4}{96}, \frac{5\pi^5}{1536}, \frac{\pi^6}{960}, \ldots,$$

**40.** Generalmente,

$$S_{2n}\left(1 - \frac{1}{2^{2n}}\right) = 1 + \frac{1}{2^{2n}} + \frac{1}{3^{2n}} + \frac{1}{4^{2n}} + \frac{1}{5^{2n}} + \cdots$$

$$- \frac{1}{2^{2n}} \qquad - \frac{1}{4^{2n}} \qquad - \cdots$$

$$= T_2 n.$$

Esto puede utilizarse para comprobar el trabajo numérico del ejemplo 38 y del ejemplo 39:

$$\frac{1}{6} \cdot \frac{3}{4} = \frac{1}{8}, \frac{1}{90} \cdot \frac{15}{16} = \frac{1}{96}, \quad \frac{1}{945} \cdot \frac{63}{64} = \frac{1}{960}.$$

**41.** $$\int_0^1 (1 - x^2)^{-1/2} \operatorname{arc\,sen} x \, . \, dx$$

$$= \int_0^1 (1 - x^2)^{-1/2} x \, dx + \frac{1}{2}\frac{1}{3} \int_0^1 (1 - x^2)^{-1/2} x^3 \, dx + \cdots$$

$$= 1 + \frac{1}{2}\frac{1}{3}\frac{2}{3} + \frac{1}{2}\frac{3}{4}\frac{1}{5}\frac{2 \cdot 4}{3 \cdot 5} + \frac{1}{2}\frac{3}{4}\frac{5}{6}\frac{1}{7}\frac{2 \cdot 4 \cdot 6}{3 \cdot 5 \cdot 7} + \cdots$$

$$= 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots$$

Ahora, evaluar la integral desde la que partimos ($=(\pi/2)^2/2$) y utilizar el ejemplo 40. Cf. EULER, *Opera Omnia*, ser. 1, vol. 14, págs. 178-181.

**42.** $$\int_0^1 (1 - x^2)^{-1/2} \operatorname{arc\,sen} x)^2 dx$$

$$= \int_0^1 (1 - x^2)^{-1/2} x^2 dx + \frac{2}{3}\frac{1}{2} \int_0^1 (1 - x^2)^{-1/2} x^4 dx + \cdots$$

$$= \frac{1}{2}\frac{\pi}{2} + \frac{2}{3}\frac{1}{2} \cdot \frac{1}{2}\frac{3}{4}\frac{\pi}{2} + \frac{2}{3}\frac{4}{5}\frac{1}{3} \cdot \frac{1}{2}\frac{3}{4}\frac{5}{6}\frac{\pi}{2} + \cdots$$

$$= \frac{\pi}{4}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots\right).$$

Evaluar, ahora, la integral de la que partimos ($=(\pi/2)^3/3$). El desarrollo de (arc sen $x$)$^2$ que hemos usado será derivada en el ejemplo 5.1. Cf. EULER, *Opera Omnia*, ser. 1, vol. 14, págs. 181-184.

**43.** (*a*) $$\sum_{n=1}^{\infty} \frac{x^n}{n^2} = \int_0^x \sum_{n=1}^{\infty} \frac{t^{n-1}}{n}\, dt = -\int_0^x t^{-1} \log(1-t)dt;$$

integrar por partes e introducir luego, como nueva variante de integración $s = 1 - t$.

(*b*) $x = 1/2$, que hace el mayor de los dos valores, $x$ y $1 - x$, tan pequeño como es posible.

**44.** Si $P_n(x) = 0$, tenemos

$$\left(1 + \frac{ix}{n}\right)^n = \left(1 - \frac{ix}{n}\right)^n,$$

$$1 + \frac{ix}{n} = e^{2\pi ki/n}\left(1 - \frac{ix}{n}\right),$$

$$x = \frac{n}{i}\,\frac{e^{\pi ki/n} - e^{-\pi ki/n}}{e^{\pi ki/n} + e^{-\pi ki/n}} = n \tan\frac{k\pi}{n},$$

donde $k = 0, 1, 2, \ldots n - 1$ si $n$ es impar.

**45.** Si $n$ es impar, podemos aceptar la expresión de las raíces; ver ejemplo 44,

$$k = 0, \pm 1, \pm 2, \ldots \pm (n-1)/2.$$

Por tanto,

$$\frac{P_n(x)}{x} = \prod_{k=1}^{(n-1)/2} \left(1 - \frac{x^2}{n^2 \tan^2(k\pi/n)}\right).$$

Obsérvese que, para $k$ fijo,

$$\lim_{n\to\infty} n \tan(k\pi/n) = k\pi.$$

Sólo se necesita un paso relativamente pequeño para llevarnos desde el punto alcanzado a una prueba que es aceptable de acuerdo con las normas modernas. Un arreglo algo diferente del argumento de Euler, debido a Cauchy, parece haber servido de modelo a Abel cuando descubrió, por analogía, la representación de las funciones elípticas en pro-

ductos infinitos. Cf. A. CAUCHY, *Oeuvres Complètes*, ser. 2, vol. 3, páginas 462-465, y N. H. ABEL, *Oeuvres Complètes*, vol. 1, págs. 335-343.

**46.** La suma de un número finito de términos es la misma cualquiera que sea el orden en que se tomen los términos. El error fue extender esto a un número infinito de términos, esto es, suponer que la suma de una *serie infinita* es la misma cualquiera que sea el orden en que se tomen los términos. Nuestro ejemplo muestra que tal declaración es falsa. La protección contra tal error es volver a las definiciones de los términos utilizados y confiar sólo en pruebas rigurosas, basadas en estas definiciones. Así, la suma de una serie infinita es, por definición, el *límite* de cierta secuencia (de la secuencia de las «sumas parciales»), e intercambiando una infinidad de términos, como hicimos, cambiamos esencialmente la secuencia definidora. (Bajo cierta *condición restrictiva* una reordenación de los términos de una serie infinita no cambia la suma; ver HARDY, *Pure Mathematics*, págs. 346-347, 374, 378-379. Sin embargo, en el presente caso no se satisface esta condición.)

Sin solución: **5, 7, 9, 10, 11, 18, 19, 20, 21.**

## CAPITULO III

**1.** Sí: $C = 2n$, $V = n + 2$, $A = 3n$.

**2.** (1) Sí: $C = m\,(p + 1)$, $V = pm + 2$, $A = m\,(p + 1) + pm$. (2) $p = 1$, $m = 4$.

**3.** (1) Excluir por un momento el tetraedro; los seis poliedros que quedan forman tres pares. Los dos poliedros del mismo par, como el cubo y el octaedro, están relacionados de modo que tienen las mismas aristas, pero el número de caras del uno es igual al número de vértices del otro. El tetraedro se queda solo, pero está relacionado consigo mismo, de forma peculiar. (2) Tomemos el cubo. Unamos los centros de todas las caras vecinas por medio de una línea recta. Las doce rectas así obtenidas forman las aristas de un octaedro regular. Este octaedro está inscrito en el cubo, sus seis vértices están en los centros de las seis caras del cubo. Recíprocamente, los centros de las ocho caras del octaedro regular son los ocho vértices de un cubo inscrito en el octaedro. Una relación recíproca semejante se da entre los poliedros del mismo par también en los otros casos. (Utilícense modelos de cartulina para el dodecaedro y el icosaedro.) El tetraedro tiene esta peculiar relación consigo mismo: los centros de sus cuatro caras son los vértices de un

tetraedro inscrito. (3) El paso de un poliedro de un par al otro mantiene la fórmula de Euler.

**4.** La esfera es dividida en $C$ países por $A$ líneas divisorias *rojas;* hay $V$ puntos que pertenecen a la línea divisoria de más de dos países. Elegir un punto en cada país: la «capital». Unir las capitales de cada dos países vecinos por una «carretera» de modo que cada carretera cruce sólo una línea divisoria y las diferentes carreteras no se crucen entre sí; dibujar estas carreteras en *azul.* Hay exactamente $A$ líneas azules (carreteras); dividen la esfera de $C'$ países con $V'$ puntos que pertenecen a la frontera de tres o más de estos países. $V' = C$ y $C' = V$. La relación entre las subdivisiones rojas y azules de la esfera es recíproca; el paso de una o otra mantiene la fórmula de Euler.

**5.** La fórmula de Euler valdrá después de haber puesto el «tejado» (sección 4) si, y sólo si, valía antes. Pero, poniendo un tejado a todas las caras no triangulares de un poliedro dado obtenemos otro poliedro con caras triangulares sólo.

**6.** Análogo al ejemplo 5: «truncar» introduce vértices con tres aristas como «poner un tejado» introduce caras triangulares. Podríamos también reducir el caso presente al ejemplo 5 utilizando el ejemplo 4.

**7.** (1) $N_0 = V, N_1 = A, N_2 = C - 1$. Los suscritos 0, 1, 2 indican la dimensionalidad respectiva; véase sección 7. (2) $N_0 - N_1 + N_2 = 1$.

**8.** (1) Establezcamos $l + m = c_1$, $lm = c_2$. Luego,

$$\begin{aligned} N_0 &= (l+1)(m+1) &&= 1 + c_1 + c_2, \\ N_1 &= (l+1)m + (m+1)l &&= c_1 + 2c_2, \\ N_2 &= lm &&= c_2. \end{aligned}$$

(2) Sí, $N_0 - N_1 + N_2 = 1$, aunque esta sencilla subdivisión de un rectángulo no puede ser generada exactamente de la misma manera descrita en el ejemplo 7.

**9.** $N_2\, 180° = (N_0 - 3)\, 360° + 180°$. Al intentar acercarnos a nuestro objetivo, que es la ecuación (2) de la solución del ejemplo 7, transformamos ésta sucesivamente en

$$\begin{aligned} 2N_0 - N_2 - 5 &= 0, \\ 2N_0 - 3N_2 + 2N_2 - 3 &= 2, \end{aligned}$$

Contando las aristas de dos maneras diferentes, obtenemos

$$3N_2 = 2N_1 - 3.$$

Estas dos ecuaciones dan como resultado

$$N_0 - N_1 + N_2 = 1,$$

que prueba la fórmula de Euler, en vista del ejemplo 7 (2).

**10.** (1) Establezcamos $l + m + n = c_1$, $lm + ln + mn = c_2$ y $lmn = c_3$. Luego,

$$\begin{aligned} N_0 &= (l+1)(m+1)(n+1) = 1 + c_1 + c_2 + c_3, \\ N_1 &= l(m+1)(n+1) + m(l+1)(n+1) + n(l+1)(m+1) \\ &= c_1 + 2c_2 + 3c_3, \\ N_2 &= (l+1)mn + (m+1)ln + (n+1)lm = c_2 + 3c_3, \\ N_3 &= lmn = c_3 \end{aligned}$$

(2) Sí, $N_0 - N_1 + N_2 - N_3 = 1$.

**11.** Tratamos con el caso $n = 3$ de la sección 16. Aquí no utilizamos ninguna circunstancia simplificadora que sería específica del caso particular $n = 3$. Por tanto, este caso particular puede «representar» muy bien el caso general (en el sentido del ejemplo 2.10) como fue indicado ya en la sección 17. El lector repetirá la discusión de la sección 16, diciendo $n$ para 3, $n + 1$ para 4, $P_n$ para 7, y $P_{n+1}$ para 11, con un poco de precaución. Ver también el ejemplo 12.

**12.** Seguir las sugerencias de la sección 17 y la analogía del ejemplo 11. Dados $n$ planos en posición general, que cortan el espacio en $S_n$ partes, añadir uno más; éste está intersectado por los planos anteriores en $n$ líneas rectas, las cuales, *estando en una posición general*, determinan sobre éste $Pn$ regiones. Cada una de estas regiones de planos actúa como un «diafragma»; divide un antiguo compartimiento del espacio (uno de aquellos $S_n$ compartimientos) en dos nuevos compartimientos, hace a uno de los antiguos desaparecer y aparecen dos nuevos compartimientos, y, finalmente, añade una unidad al anterior número $S_n$ de compartimientos. De aquí la relación que deseamos probar.

**13.** Véase la tercera columna de la tabla en la sección 14.

**14.** La segunda columna de la tabla de la sección 14 concuerda con

$$S_n = 1 + n + \frac{n(n-1)}{1 \cdot 2} + \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3}$$

$$= \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3};$$

utilizamos la notación usual para coeficientes de binomios.

**15.** Tres finitas, ocho infinitas.

**16.** $P_n^\infty$ denota el número de aquellas de entre las $P_n$ partes definidas en el ejemplo 11 que son infinitas. Por observación, para

$$n = 1, 2, 3$$
$$P_n^\infty = 2, 4, 6.$$

Intuir: $P_n^\infty = 2n$. Prueba: Tomar un punto en una de las partes finitas e imaginar un círculo siempre creciente con este punto como centro. Cuando este círculo llega a ser muy grande, las $P_n - P_n^\infty$ partes finitas coinciden prácticamente con su centro. Ahora, $n$ líneas diferentes, pasando por el centro del círculo, cortan la periferia en $2n$ puntos y la dividen en $2n$ partes. De aquí que, en efecto, $P_n^\infty = 2n$.

$$P_n - P_n^\infty = 1 - n + \frac{n(n-1)}{2}.$$

Por ejemplo, la respuesta al ejemplo 15 es

$$1 - 4 + 6 = 3.$$

**17.** Lo mismo que en el ejemplo 18, por analogía con la solución del ejemplo 16.

**18.** Lo mismo que en el ejemplo 19.

**19.** Véase el ejemplo 20.

**20.** Consideremos $n$ círculos en el plano; dos cualesquiera de ellos se intersectan en posición general. Llamamos $S_n^\infty$ al número de partes en que estos círculos dividen al plano, en vista de los ejemplos 17, 18 y 19. En analogía con la sección 16 obsérvese que el número de partes en que es dividido un círculo por $n$ círculos que se cortan entre sí es $2n$ (se supone

que están en posición general). Obsérvese (pensar en las tres interpretaciones de $S_n^\infty$):

$$\begin{aligned} n &= 1 \quad 2 \quad 3 \quad 4 \\ 2n &= 2 \quad 4 \quad 6 \quad 8 \\ S_n^\infty &= 2 \quad 4 \quad 8 \quad 14. \end{aligned}$$

Intuición: $S_{n+1}^\infty = S_n^\infty + 2n$. Prueba: como en los ejemplos 11 y 12. Por ejemplo,

$$S_5^\infty = S_4^\infty + 8 = 14 + 8 = 22;$$

Esta es la solución de los ejemplos 17, 18 y 19. Otras intuiciones:

$$S_n^\infty = 2\binom{n}{0} + 2\binom{n}{2},$$

$$S_n - S_n^\infty = -\binom{n}{0} + \binom{n}{1} - \binom{n}{2} + \binom{n}{3}.$$

**21.** Ver ejemplos 22-30.

**22.** Equivocado: $C = 1, V = A = 0, 1 + 0 \neq 0 + 2.$

**23.** Equivocado: $C = 2, V = 0, A = 1, 2 + 0 \neq 1 + 2.$

**24.** Equivocado: $C = 3, V = 0, A = 2, 3 + 0 \neq 2 + 2.$

**25.** Correcto: $C = 3, V = 2, A = 3, 3 + 2 = 3 + 2.$

**26.** Equivocado: $C = p + 1, V = 0, A = p, (p + 1) + 0 \neq p + 2$; ver los ejemplos 22, 23 y 24 para los casos $p = 0, 1, 2$, respectivamente. Obsérvese que, en el presente caso, la solución del ejemplo 2 (1) resulta inaplicable.

**27.** El caso $m = 3, p = 0$ es correcto; ver el ejemplo 25; y así es, más generalmente, el caso $m \geqq 3 : C = m, V = 2, A = m, m + 2 = m + 2$. El caso $m = 0, p = 0$ está equivocado; ver el ejemplo 22. Los dos casos restantes *pueden* ser interpretados de modo que parezcan correctos. (1) $m = 1, p = 0$: un país con una barrera interior que tiene dos extremos, $C = 1, V = 2, A = 1, 1 + 2 = 1 + 2$. (2) $m = 2, p = 0$: dos países separados por dos arcos y dos ángulos, $C = 2, V = 2, A = 2, 2 + 2 = 2 + 2$. La interpretación más obvia, dada en el ejemplo 23, resulta «errónea». Con la interpretación presente la solución del ejemplo 2 (1) permanece inaplicable al caso $m > 0, p = 0$.

**28.** $m \geqq 3, p \geqq 1$. La prueba utiliza el hecho de que, en cualquier poliedro convexo, al menos tres aristas rodean una cara y también por lo menos tres aristas se encuentran en un vértice.

**29.** Los ejemplos 22-28 sugieren dos condiciones: (1) Un país, considerado en $C$ como una cara de un poliedro convexo, sería del «tipo de una región circular»; una esfera completa no es de este tipo, ni tampoco un anillo. (2) Una línea considerada en $A$, como una arista de un poliedro convexo, terminaría en ángulos; la periferia completa de un círculo no termina así (no tiene terminación). El ejemplo 22 no satisface a (1), ni el 23 a (2); el ejemplo 24 no satisface a ninguno de los dos y el ejemplo 25 o, más generalmente, el caso $m > 0, p = 0$, interpretado como en la solución del ejemplo 27, satisface a (1) y (2).

**30.** (1) Tomar el caso (3, 2) del ejemplo 2 (1); cf. el ejemplo 26' pero borrar en cada meridiano el arco entre los dos círculos paralelos: $C = 7, V = 8, A = 12, 7 + 8 \neq 12 + 2$, hay una zona esférica entre los países y, por tanto, están en conflicto con la condición (1), pero no con la (2), del ejemplo 29. (2) $C = 1, V = 1, A = 0$ (un país, que cubriera todo el globo, excepto un punto matemático en el polo Norte); correcto, $1 + 1 = 0 + 2$, no hay conflicto con (1) o (2) del ejemplo 29; etc.

**32.** $3C_3 + 4C_4 + 5C_5 + \ldots = 3V_3 + 4V_4 + 5V_5 + \ldots = 2A.$

**33.** $4\pi$, $12\ \pi$, $8\pi$, $36\pi$, $20\pi$, ..., respectivamente.

**34.** $\Sigma\alpha = \pi C_3 + 2\pi C_4 + 3\pi C_5 + \ldots$

**35.** Por los ejemplos 34, 32, 31,

$$\Sigma\alpha = \pi\,\Sigma(n-2)C_n = 2\pi(A - C).$$

**36.** Un polígono esférico convexo con $n$ lados puede ser dividido en $n - 2$ triángulos esféricos. Por tanto,

$$\begin{aligned} A &= \alpha_1 + \alpha_2 + \ldots + \alpha_n - (n-2)\pi \\ &= 2\pi - (\pi - \alpha_1) - (\pi - \alpha_2) - \ldots - (\pi - \alpha_n) \\ &= 2\pi - a_1' - a_2' - \ldots - a_n' \\ &= 2\pi - P'. \end{aligned}$$

**37.** Las caras de los poliedros que pasan por uno de los vértices comprenden un ángulo interior sólido; su suplemento fue llamado por Descartes al ángulo exterior sólido. Describir una esfera con radio 1 en torno al vértice, pero coger sólo el sector de la esfera contenido en el

ángulo sólido exterior; los sectores generados de este modo en los distintos vértices del poliedro forman, al ser agrupados, una esfera completa, como los sectores circulares de la figura plana análoga (figura 3.7) forman, agrupados, un círculo completo. Consideramos como medida de un ángulo sólido el área del correspondiente polígono esférico: en total, la medida de todos los ángulos exteriores sólidos del poliedro es, en efecto, $4\pi$.

**38.** $P_1, P_2, \ldots P_V$ denotan los perímetros de los polígonos esféricos que corresponden a los ángulos sólidos interiores $V$ del poliedro. Luego, por los ejemplos 36 y 37,

$$\begin{aligned}\Sigma\alpha &= P_1 + P_2 + \ldots + P_V \\ &= 2\pi - A_1' + 2\pi - A_2' + \ldots + 2\pi - A_V' \\ &= 2\pi V - 4\pi.\end{aligned}$$

**39.** Por los ejemplos 35 y 38,

$$2\pi(A - C) = \Sigma\alpha = 2\pi(V - 2).$$

**40.** Por los ejemplos 31 y 32,

$$\begin{aligned}3C &= 3C_3 + 3C_4 + 3C_5 + \ldots \\ &\leqq 3C_3 + 4C_4 + 5C_5 + \ldots = 2A,\end{aligned}$$

que nos dan las primeras de las seis desigualdades propuestas. La igualdad se alcanza sólo cuando $C = C_3$, es decir, cuando todas las caras son triángulos. Eliminando $A$ y luego $C$ a partir del teorema de Euler y la desigualdad probada obtenemos las dos desigualdades que quedan en la primera fila; éstas pasan a ser ecuaciones si, y sólo si, todas las caras son triángulos. Intercambiando los papeles de $C$ y $V$, como sugieren los ejemplos 3 y 4, obtenemos las tres desiguadades propuestas en la segunda fila; éstas pasan a ser ecuaciones si, y sólo si, todos los vértices del poliepro tienen tres aristas. Algunas de las desigualdades probadas las encontramos en las notas de Descartes.

**41.** Por el teorema de Euler,

$$6C - 2A = 12 + 2\,(2A - 3\,V),$$

y de aquí, por los ejemplos 31, 32 y 40.

$$\begin{aligned}3C_3 + 2C_4 + C_5 &= 12 + 2(2A - 3V) + C_7 + 2C_8 + \ldots \\ 3C_3 + 2C_4 + C_5 &\geqq 12,\end{aligned}$$

y así, *todos los poliedros convexo sdeben tener algunas caras con menos de seis lados.*

Sin solución: **31.**

## CAPITULO IV

**1.** $R_2(25) = 12$; ver sección 2; $S_3(11) = 3$.

**2.** $R_2(n)$ denota el número de los puntos de cruce en un plano que está sobre la periferia de un círculo con radio $\sqrt{n}$ y centro en el origen. (Tomemos el caso $n = 25$, ejemplo 1, y dibujemos el círculo.) $R_3(n)$ es el número de los puntos d ecruce en el espacio sobrela superficie de la esfera con radio $\sqrt{n}$ y centro en el origen.

**3.** Si $p$ es un primo impar, $R_2(p^2) = 12$ ó 4, según que $p$ dividido por 4 deje 1 ó 3 en el resto.

**4.** La comparación de las tablas sugiere: si $p$ es un primo impar, tanto $p$ como $p^2$ son expresables como la suma de dos cuadrados, o bien ninguno de los dos lo son. Una conjetura más precisa es también apoyada por nuestras observaciones: si $p$ es un primo impar, $R_2(p) = 8$ ó 0, según que $p$ dividido por 4 dé 1 ó 3 en el resto.

**5.** Si $p = x^2 + y^2$, se sigue que

$$p^2 = x^4 + 2x^2y^2 + y^4 = (x^2 - y^2)^2 + (2xy)^2.$$

Es decir, si $R_2(p) > 0$, también $R_2(p^2) > 0$. Esto es sólo la mitad de nuestra conjetura menos precisa, y una parte pequeña de la más precisa. (Si sabemos que $R_2(p^2) > 0$, una conclusión referente a $R_2(p)$ es, sin duda, menos obvia.) Aun así, parece que tal verificación parcial fortalece mucho nuestra confianza en la conjetura menos precisa y también algo la más precisa.

**6.** $R_3(p) = 0$ para $n = 7, 15, 23$ y 28, para ninguna otra $n$ superior a 30; véase la tabla II.

**7.** Las respectivas contribuciones a $S_4(n)$ son: (1) 24, (2) 12, (3) 6, (4) 4, (5) 1.

**8.** Primero, referirse a los casos señalados en el ejemplo 7. Si $S_4(4u)$ es impar, el caso (5) surge necesariamente, y así,

$$4u = a^2 + a^2 + a^2 + a^2,$$

$$u = a^2.$$

Segundo, a todo divisor $d$ de $u$ corresponde el divisor $u/d$ y estos dos divisores son diferentes a menos que $u = d^2$. En consecuencia, el número de divisores de $u$ es par o impar, según que $u$ sea o no un cuadrado, y lo mismo vale para la suma de estos divisores, puesto que cada divisor de $u$ es impar, como $u$ mismo. Nosotros conjeturamos en la sección 6 que $S_4(4u)$ y la suma de los divisores de $u$ son iguales; hemos probado ahora que estos dos números dan el mismo resto cuando se dividen por 2. Habiendo probado una parte de nuestra conjetura, tenemos, por supuesto, más fe en ella.

**9.** (1) $24 \times 2^4 = 8 \times 48$ (6) $24 \times 2^3 = 8 \times 24$
(2) $12 \times 2^4 = 8 \times 24$ (7) $12 \times 2^3 = 8 \times 12$
(3) $6 \times 2^4 = 8 \times 12$ (8) $4 \times 2^3 = 8 \times 4$
(4) $4 \times 2^4 = 8 \times 8$ (9) $12 \times 2^2 = 8 \times 6$
(5) $1 \times 2^4 = 8 \times 2$ (10) $6 \times 2^2 = 8 \times 3$
(11) $4 \times 2^2 = 8 \times 1$.

**10.** Ver la tabla II, págs. . Comprobar, por lo menos, algunas entradas. Se sigue del ejemplo 9 que $R_4(n)$ es divisible por 8.

**11.** Intentando observar algunas regularidades fragmentarias (como hicimos en la sección 6) podemos agrupar algunos de los casos más notables como sigue:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) | 2 | 3 | 5 | 7 | 11 | 13 | 17 | 19 | 23 | 29 |
| | 3 | 4 | 6 | 8 | 12 | 14 | 18 | 20 | 24 | 30 |
| (2) | | 2 | | 4 | | 8 | | 16 | | |
| | | 3 | | 3 | | 3 | | 3 | | |
| (3) | 4 | 8 | 12 | 16 | 20 | 24 | 28 | | | |
| | 3 | 3 | 12 | 3 | 18 | 12 | 24. | | | |

En (1), (2) y (3) la primera línea da $n$, la segunda $R_4(n)$.

**12.** Hecho en la solución del ejemplo 11: (1) primos, (2) valores de 2, (3) números divisibles por 4.

**13.** Por analogía con la sección 6 y un poco de observación, la ley es relativamente fácil de descubrir cuando $n$ no es divisible por 4. En consecuencia, nos concentramos sobre el caso (3) en la solución del ejemplo 11.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $n$ | = | 4 | 8 | 12 | 16 | 20 | 24 | 28 |
| $n/4$ | = | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $R_4(n)/8$ | = | 3 | 3 | 12 | 3 | 18 | 12 | 24. |

Los números en tipos gruesos de imprenta de la tercera línea son la *suma de todos los divisores* del número correspondiente en la segunda línea —y, por tanto, la suma de *algunos* divisores del número correspondiente en la primera línea, en la cual estamos realmente interesados. Esta observación nos conduce a otro intento:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $n$ | = | 4 | 8 | 12 | | 16 | |
| $R_4(n)/8$ | = | $1+2$ | $1+2$ | $1+2+3+6$ | | $1+2$ | |
| $n$ | = | 20 | | 24 | | 28 | |
| $R_4(n)/8$ | = | $1+2+5+10$ | | $1+2+3+6$ | | $1+2+7+14.$ | |

¿Qué divisores son sumados juntos? ¿Cuáles son omitidos?

**14.** $R_4(n)$, el número de representaciones de $n$ como una suma de cuatro cuadrados, es igual a ocho veces la suma de esos divisores de $n$ que no son divisibles por 4. (Si $n$ misma no es divisible por 4, ninguno de sus divisores lo es, y, por ello, la regla es más frecuente en este caso más sencillo.)

**15.** De acuerdo con las columnas de la tabla II:

| | | | |
|---|---|---|---|
| 31 | $25+4+1+1$ | $12\times 16$ | $32=31+1$ |
| | $9+9+9+4$ | $4\times 16$ | |
| 32 | $16+16$ | $6\times 4$ | $3=2+1$ |
| 33 | $25+4+4$ | $12\times 8$ | $48=33+11+3+1$ |
| | $16+16+1$ | $12\times 8$ | |
| | $16+9+4+4$ | $12\times 16$ | |

**16.** $5=1+1+1+1+1=4+1$

$$R_8(5)=\binom{8}{5}2^5+8\cdot 7\cdot 2^2=2016=16\times 126.$$

$$40=25+9+1+1+1+1+1+1$$
$$40=9+9+9+9+1+1+1+1$$
$$S_8(40)=8\cdot 7+\binom{8}{4}=126.$$

**17.** Ejemplo 16. La tabla III ha sido construidda realmente por un método menos laborioso que la del ejemplo 16; cf. ejemplos 6.17 y 6.23.

**18.** Dentro de los límites de la tabla III, tanto $R_8(n)$ como $S_8(8n)$ aumentan continuamente con $n$, mientras $R_4(n)$ y $S_4(4(2n-1))$ oscilan irregularmente.

**19.** La analogía con $R_4(n)$ y $S_4(4(2n-1))$ apunta a los divisores. Una regularidad fragmentaria es fácil de observar: si $n$ es impar, $R_8(n)/16$ y $S_8(8n)$ son exactamente iguales; si $n$ es par, son diferentes, aunque la diferencia es relativamente pequeña en la mayoría de los casos.

**20.** Pares e impares ya en el ejemplo 19. Valores de 2:

| $n$ | 1 | 2 | 4 | 8 | 16 |
|---|---|---|---|---|---|
| $S(8n)$ | 1 | 8 | 64 | 512 | 4096. |

También la segunda línea consta de valores de 2:

| $n$ | $2^0$ | $2^1$ | $2^2$ | $2^3$ | $2^4$ |
|---|---|---|---|---|---|
| $S(8n)$ | $2^0$ | $2^3$ | $2^6$ | $2^9$ | $2^{12}$. |

¿Cuál es la ley de los exponentes?

**21.** Si $n$ es un valor de 2, $S(8n) = n^3$. Esto (y el ligero incremento de $R_8(n)$ y $S_8(8n)$) conduce a construir la siguiente tabla:

| $n$ | $R_8(n)/16 - n^3$ | $S_8(8n) - n^3$ |
|---|---|---|
| 1 | 0 | 0 |
| 2 | —1 | 0 |
| 3 | 1 | 1 |
| 4 | 7 | 0 |
| 5 | 1 | 1 |
| 6 | —20 | 8 |
| 7 | 1 | 1 |
| 8 | 71 | 0 |
| 9 | 28 | 28 |
| 10 | —118 | 8 |
| 11 | 1 | 1 |
| 12 | 260 | 64 |
| 13 | 1 | 1 |
| 14 | —336 | 8 |
| 15 | 153 | 153 |
| 16 | 583 | 0 |
| 17 | 1 | 1 |
| 18 | —533 | 224 |
| 19 | 1 | 1 |
| 20 | 946 | 64 |

En la columna que se refiere a $R_8(n)$ los signos + y — son distribuidos regularmente.

**22.** ¡Cubos de divisores!

| $n$ | $R_8(n)/16 = S_8(8n)$ |
|---|---|
| 1 | $1^3$ |
| 3 | $3^3 + 1^3$ |
| 5 | $5^3 + 1^3$ |
| 7 | $7^3 + 1^3$ |
| 9 | $9^3 + 3^3 + 1^3$ |
| 11 | $11^3 + 1^3$ |
| 13 | $13^3 + 1^3$ |
| 15 | $15^3 + 5^3 + 3^3 + 1^3$ |
| 17 | $17^3 + 1^3$ |
| 19 | $19^3 + 1^3$ |

| $n$ | $R_8(n)/16$ | $S_8(8n)$ |
|---|---|---|
| 2 | $2^3 - 1^3$ | $2^3$ |
| 4 | $4^3 + 2^3 - 1^3$ | $4^3$ |
| 6 | $6^3 - 3^3 + 2^3 - 1^3$ | $6^3 + 2^3$ |
| 8 | $8^3 + 4^3 + 2^3 - 1^3$ | $8^3$ |
| 10 | $10^3 - 5^3 + 2^3 - 1^3$ | $10^3 + 2^3$ |
| 12 | $12^3 + 6^3 + 4^3 - 3^3 + 2^3 - 1^3$ | $12^3 - 4^3$ |
| 14 | $14^3 - 7^3 + 2^3 - 1^3$ | $14^3 + 2^3$ |
| 16 | $16^3 + 8^3 + 4^3 + 2^3 - 1^3$ | $16^3$ |
| 18 | $18_8 - 9^3 + 6^3 - 3^3 + 2^3 - 1^3$ | $18^3 + 6^3 + 2^3$ |
| 20 | $20^3 + 10^3 - 5^3 + 4^3 + 2^3 - 1^3$ | $20^3 + 4^3$ |

**23.** (1) $(-1)^{n-1}\, R_8(n)/16$ es igual a la suma de los cubos de todos los divisores impares de $n$, menos la suma de los cubos de todos los divisores pares de $n$. (2) $S_8(8n)$ es igual a la suma de los cubos de aquellos divisores de $n$ cuyos co-divisores son impares. (Si $d$ es un divisor de $n$, llamamos $n/d$ a los co-divisores de $d$.) Ver ejemplo 6.24 para la historia de estos teoremas y referencias.

**24.** Construir la tabla

| | | | | |
|---|---|---|---|---|
| 0 | 3 | 6 | 9 | 12 |
| 5 | 8 | 11 | 14 | |
| 10 | 13 | | | |

que imaginaremos se extiende sin límite hacia la derecha y hacia abajo, y mostrar que los únicos enteros positivos no expresables en la forma propuesta son 1, 2, 4, 7,

**25.** Caso $a = 3$, $b = 5$ en el ejemplo 24; $a$ y $b$ son co-primos. El último entero no expresable $ab - a - b = (a - 1)(b - 1) - 1$. Esto es incomparablemente más fácil de probar que las leyes referentes a las sumas de los cuadrados.

**26.** (1) es generalmente cierto. (2) no es generalmente cierto, pero la primera excepción es $n = 341$. (Ver G. H. HARDY y E. M. WRIGHT, *An introduction to the theory of numbers*, Oxford, 1938, págs. 69, 72.)

## CAPITULO V

**1.** [Cf. Putnam, 1948]

(*a*) $$x + \frac{2}{3}x^3 + \frac{2}{3}\frac{4}{5}x^5 + \ldots + \frac{2}{3}\frac{4}{5}\frac{6}{7}\cdots\frac{2n}{2n+1}x^{2n+1} + \ldots$$

(*b*) Habiendo verificado la ecuación diferencial pongamos

$$y = a_0x + a_1x^3 + a_2x^5 + \ldots + a_nx^{2n+1} + \ldots.$$

Para comparar los coeficientes de valores iguales podemos usar el cuadro

| | 1 | $x^2$ | $x^4$ | ... | $x^{2n}$ | |
|---|---|---|---|---|---|---|
| $y'$ | $a_0$ | $3a_1$ | $5a_2$ | ... | $(2n+1)a_n$ | ... |
| $-x^2y'$ | | $-a_0$ | $-3a_1$ | ... | $-(2n-1)_{n-1}$ | ... |
| $-xy$ | | $-a_0$ | $-a_1$ | ... | $-a_{n-1}$ | ... |
| | 1 | 0 | 0 | ... | 0 | ... |

que da $a_0 = 1$ y $(2n + 1)a_n = 2na_{n-1}$ para $n \geqq 1$.

**2.** [Cf. Putnam, 1950]

$$(a)\quad y = \frac{x}{1} + \frac{x^3}{1\cdot 3} + \frac{x^5}{1\cdot 3\cdot 5} + \cdots + \frac{x^{2n-1}}{1\cdot 3\cdot 5 \ldots (2n-1)} + \cdots$$

(*b*) Este desarrollo satisface

$$y' = 1 + \frac{x^2}{1} + \frac{x^4}{1\cdot 3} + \frac{x^6}{1\cdot 3\cdot 5} + \cdots,$$

$$y' = 1 + xy.$$

El producto dado $y$ satisface la misma ecuación diferencial. Tanto la extensión como el producto se desvanecen cuando $x = 0$. De aquí que sean idénticos.

**3.** Las relaciones entre los coeficientes $a_n$ derivados de

$$\frac{1}{1+x} + \frac{4a_1x}{(1+x)^3} + \frac{16a_2x^2}{(1+x)^5} + \cdots$$

$$= 1 + a_1x^2 + a_2x^4 + \cdots = f(x)$$

son mostrados en el cuadro (ver ejemplo 1)

| | | | | |
|---|---|---|---|---|
| 1 | —1 | 1 | —1 | 1 |
| | $4a_1$ | $-4a_1\cdot 3$ | $4a_1\cdot 6$ | $-4a_1\cdot 10$ |
| | | $16a_2$ | $-16a_2\cdot 5$ | $16a_2\cdot 15$ |
| | | | $64a_3$ | $-64a_3\cdot 7$ |
| | | | | $256a_4$ |

---

| | | | | |
|---|---|---|---|---|
| 1 | 0 | $a_1$ | 0 | $a_2$ |

Estos dan

$$f(x) = 1 + \left(\frac{1}{2}\right)^2 x^2 + \left(\frac{1}{2}\frac{3}{4}\right)^2 x^4 + \left(\frac{1}{2}\frac{3}{4}\frac{5}{6}\right)^2 x^6 + \left(\frac{1}{2}\frac{3}{4}\frac{5}{6}\frac{7}{8}\right)^2 x^8 + \cdots$$

Este ejemplo es de interés histórico. Ver GAUSS, *Werke*, vol. 3, páginas. 365-369.

**4.** Estudiar la ordenación del siguiente cuadro (ver ejemplo 1, 3):

| | | | | | |
|---|---|---|---|---|---|
| $f(x)^3$ | $a_0^3$ | $3a_1a_0^2$ | $3a_2a_0^2$ | $3a_3a_0^2$ | $3a_4a_0^2$ |
| | | | $3a_1a_0$ | $6a_2a_1a_0$ | $6a_3a_1a_0$ |
| | | | | $a_1^3$ | $3a_2^2a_0$ |
| | | | | | $3a_2a_1^2$ |
| $f(x)f(x^2)$ | $3a_0^2$ | $3a_1a_0$ | $3a_2a_0$ | $3a_3a_0$ | $3a_4a_0$ |
| | | | $3a^\circ a_1$ | $3a_1a_1$ | $3a_2a_1$ |
| | | | | | $3a_0a_2$ |
| $2f(x^3)$ | $2a_0$ | | | $2a_1$ | |
| | $6a_1$ | $6a_2$ | $6a_3$ | $6a_4$ | $6a_5$ |

Empezando desde $a_0 = 1$ obtenemos recurrentemente $a_1$, $a_2$, $a_3$, $a_4$, y $a_5 = 8$. Ver G. POLYA, *Zeitschrift für Kristallographie*, vol. (A) 93, (1936), págs. 415-443, y *Acta Mathematica*, vol. 68 (1937), págs. 145-252.

**5.** Por comparación de los desarrollos de la sección 1.

**6.** $(a)\varepsilon^2/15$. $(b) + \infty$. En ambos casos extremos el error es positivo, el valor aproximado más grande que el valor cierto.

**7.** $(a)\varepsilon^2/15$. $(b)$ 1/3. En ambos casos extremos el valor aproximado es más grande que el valor cierto.

**8.** $4\pi(a^2 + b^2 + c^2)/3$. Hay razones para sospechar que esta aproximación da valores por exceso de los valores ciertos. Ver G. POLYA, *Publicaciones del Instituto de Matemática*, Rosario, vol. 5 (1943).

**9.** Al pasar de las integrales a series utilizamos el desarrollo binomial y las fórmulas integrales del ejemplo 2.42.

$$P = 2\pi a\left[1 - \frac{1}{2}\sum_1^\infty \frac{1}{2}\frac{3}{4}\cdots\frac{2n-1}{2n}\frac{\varepsilon^{2n}}{2n-1}\right],$$

$$P' = 2\pi a\left[1 - \sum_1^\infty \frac{3}{4}\frac{7}{8}\cdots\frac{4n-1}{4n}\frac{\varepsilon^{2n}}{4n-1}\right].$$

**10.** Utilizamos la solución del ejemplo 9 y ponemos $\frac{1}{2}\frac{3}{4}\frac{5}{6}\cdots\frac{2-1}{2n} = g_n$. Luego $g_1 > gn$ para $n \geqq 2$ y, para $\varepsilon > 0$,

$$E - P = 2\pi a\sum_2^\infty (g_{1n}\, g_n - g_n^2)\,\varepsilon^{2n}/(2n-1) > 0.$$

**11.** El término inicial del error relativo de $P''$ es

$$-[\alpha + 3(1-\alpha)]\,\varepsilon^4/64 + \ldots$$

y así es de orden 4, a menos que $\alpha = 3/2$ y $P'' = P + (P - P')/2$.

**12.** $(P'' - E)/E = 3 \cdot 2^{-11}\varepsilon^8 + \ldots$ cuando $\varepsilon$ es pequeña.

$= (3\pi - 8)/8 = .1781$ cuando $\varepsilon = 1$.

$= .00019$ cuando $\varepsilon = 4/5$.

De aquí la conjetura $P'' > E$. Ver G. PEANO, *Applicazioni geometriche del calcolo infinitesimale*, págs. 231-236.

**13.**
$$e^p = \lim_{n\to\infty}\left(1 + \frac{p}{n}\right)^n.$$

En consecuencia, la conclusión deseada es equivalente a

$$\limsup_{n\to\infty}\left(\frac{n(a_1 + a_{n+p})}{(n+p)a_n}\right)^n \geqq 1.$$

La asunción opuesta implica

$$\frac{n(a_1 + a_{n+p})}{(n+p)a_n} < 1,$$

para $n > N$, $N$ fijo o, lo que es lo mismo,

$$\frac{a_{n+p}}{n+p} - \frac{a_n}{n} < -\frac{a_1}{n+p}.$$

Considerar los valores $n = (m-1)p$:

$$\frac{a_{mp}}{mp} - \frac{a_{(m-1)p}}{(m-1)p} < -\frac{a_1}{mp},$$

$$\frac{a_{(m-1)p}}{(m-1)p} - \frac{a_{(m-2)p}}{(m-2)p} < -\frac{a_1}{(m-1)p},$$

. . . . . .

Como en la sección 5, concluimos que, con una constante $C$ adecuada,

$$\frac{a_{mp}}{m} < C - a_1\left(1 + \frac{1}{2} + \ldots + \frac{1}{m}\right)$$

y esto conduce para $m \to \infty$ a una contradicción de la hipótesis $a_n > 0$.

**14.** El ejemplo $a_n = n^c$ de la sección 4 sugiere

$$a_1 = 1,$$
$$a_n = n \log n \quad \text{para} \quad n = 2, 3, 4, \ldots$$

Con esta elección

$$\left(\frac{a_1 + a_{n+p}}{a_n}\right) = \left\{\frac{1 + (n+p)\,[\log n + \log(1 + (p/n))]}{n \log n}\right\}^n$$

$$= \left\{\frac{(n+p)\log n + 1 + (n+p)\left[\dfrac{p}{n} - \dfrac{p^2}{2n^2} + \ldots\right]}{n \log n}\right\}^n$$

$$= \left\{1 + \frac{p + \alpha_n}{n}\right\}^n \to e^p$$

puesto que $\alpha_n \to 0$.

**15.** Las mantisas en cuestión son las novecientas ordenadas de la curva que surge lentamente $y = \log x - 2$ que corresponden a las abscisas $x = 100, 101, \ldots, 999$; log significa logaritmo vulgar. La tabla I especifica cuántos de los novecientos puntos de la curva caen en una cierta lista horizontal de 1/10 de ancho. Consideremos los puntos en que la curva entra y sale de una lista semejante. Si $x_n$ es la abscisa de un punto tal

$$\log x_n - 2 = n/10,$$
$$x_n = 100 \cdot 10^{n/10},$$

donde $n = 0, 1, 2, \ldots, 10$. El número de enteros en cualquier intervalo es, aproximadamente, igual a la longitud del intervalo; la diferencia es menos de una unidad. Por tanto, el número de las mantisas en cuestión con la primera cifra $n$ es $x_{n+1} - x_n$, con un error menor a 1. Ahora,

$$x_{n+1} - x_n = 100\,(10^{1/10} - 1)\,10^{n/10}$$

es el $n$-avo término de una progresión geométrica con radio

$$10^{1/10} = 1.25893.$$

Predecir y observar el fenómeno análogo en una tabla de logaritmos vulgares de seis cifras.

**16.** La repetición periódica puede (y debe) ser considerada como una clase de simetría; pero ésta está presente en todos los casos y por

ello no la mencionaremos de nuevo. Las siguientes clases de simetría juegan un papel en nuestra clasificación.

(1) Centro de simetría. Notación: $c$, $c'$.

(2) Línea de simetría. Notación: $h$ es la línea horizontal, $v$ o $v'$ si es la vertical.

(3) Simetría deslizante: los frisos se deslizan horizontalmente y se reflejan en la línea horizontal del centro *simultáneamente*, coinciden consigo mismo (en frisos 5, 7, $a$, $b$). Notación: $g$.

Los siguientes tipos de simetría son representados en la figura 5.2. (La tilde ′ es usada para distinguir dos elementos de simetría de la misma clase, como $c$ y $c'$, o $v$ y $v'$, cuando su situación en la figura es esencialmente diferente.)

1, d: sin simetría (excepto periodicidad)

2, g: $c$, $c'$, $c$, $c'$, ...

3, f: $v$, $v'$, $v$, $v'$, ...

4, e: $h$

5, a: $g$

6, c: $h$; $(v, c)$, $(v', c')$, $(v, c)$, $v'$, $c')$, ...

7, b: $g$; $v$, $c$, $v$, $c$, ...

Todas las clases posibles de simetría son representadas en la figura 5.2, como puede usted darse cuenta de ello inductivamente.

**17.** Tres clases diferentes de simetría son representadas: 1 es del mismo tipo que 2, 3 y 4. Intentemos encontrar todos los tipos. Cf. G. Polya, *Zeitschrift für Kristallographie*, vol. 60 (1924), págs. 278-282; P. Niggli, *ibíd.*, págs. 283-298, y H. Weyl, *Symmetry*, Princeton, 1962.

**18.** No tengamos en cuenta los detalles que se desprenden de los tipos de imprenta. Luego, (1) línea vertical de simetría, (2) línea horizontal de simetría, (3) centro de simetría, (4) las tres clases precedentes de simetría unidas, (5) sin simetría. Lo mismo es válido para las cinco curvas que representan las cinco ecuaciones de las coordenadas rectangulares. Algunas variantes de este problema pueden usarse para animar una clase de geometría analítica.

## CAPITULO VI

**2.** $x(1-x)^{-2}$. Caso particular del ejemplo 3, con $f(x) = (1-x)^{-1}$; obtenerlo también por combinación de los ejemplos 4 y 5.

**3.** $xf'(x) = \sum_{n=0}^{\infty} na_n x^n$.

**4.** $xf(x) = \sum_{n=1}^{\infty} a_{n-1} x^n$.

**5.** $(1-x)^{-1} f(x) = \sum_{n=0}^{\infty} (a_0 + a_1 + \ldots + a_n) x^n$; caso particular del ejemplo 6.

**6.** $f(x)g(x) = \sum_{n=0}^{\infty} (a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0) x^n$.

**7.** $D_3 = 1$, $D_4 = 2$, $D_5 = 5$, $D_6 = 14$. Para $D_6$, referirse a la figura 6.1; hay dos disecciones diferentes del tipo I, seis del tipo II y seis del tipo III.

**8.** La fórmula de recursión es verificada para $n = 6$:

$$14 = 1 \times 5 + 1 \times 2 + 2 \times 1 + 5 \times 1.$$

Elegimos un cierto lado de la base del polígono (el lado horizontal en la figura 6.2) y empezamos la disección dibujando el segundo y el tercer lado del triángulo cuyo primer lado es la base. Habiendo elegido $\Delta$, aún tenemos que diseccionar un polígono de $k$ lados a la izquierda de $\Delta$ y otro polígono con $n + 1 - k$ lados a la derecha; ambos polígonos unidos dan $D_k D_{n+1-k}$ posibilidades. Elegimos $k = 2, 3, 4, \ldots n-1$; por supuesto, hay que interpretar adecuadamente el caso $k - 2$.

**9.** Por los ejemplos 4 y 6, la fórmula periódica para $D_n$ da

$$xg(x) = D_2 x^3 + [g(x)]^2.$$

Elegimos la solución de esta ecuación cuadrática, cuya extensión empieza con $x^2$:

$$g(x) = (x/2)\,1 - (1-4x)^{1/2}].$$

Desarrollando y usando la notación para coeficientes binomiales, encontramos:

$$D_n = -\frac{1}{2}\binom{1/2}{n-1}(-4)^{n-1}.$$

**10.** Mejor así

$$\sum_{u=-\infty}^{\infty} x^{u^2} \sum_{v=-\infty}^{\infty} x^{v^2} \sum_{w=-\infty}^{\infty} x^{w^2} = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} x^{u^2+v^2+w^2},$$

donde $u$, $v$ y $w$ alinean todos los enteros (de $-\infty$ a $+\infty$) independientemente, de modo que la triple suma se extiende sobre todos los puntos de cruce del espacio (ver el ejemplo 4.2). Para ver que ésta es la función generatriz de $R_3(n)$ recogemos los términos cuyos exponente $u^2 + v^2 + w^2$ tienen el mismo valor $n$.

**11.**
$$\sum_{n=0}^{\infty} R_k(n)x^n = \left[\sum_{n=0}^{\infty} R_1(n)x^n\right]^k.$$

**12.**
$$\sum_{n=1}^{\infty} S_k(n)x^n = \left[\sum_{n=1}^{\infty} x^{(2n-1)2}\right]^k.$$

**13.** $I$, $J$, $K$ y $L$ denotan ciertas series de valores todos cuyos coeficientes son enteros. Luego, las funciones generatrices de $R_1(n)$, $R_2(n)$, $R_4(n)$ y $R_8(n)$ son de la forma

$$\begin{aligned} &1 + 2I, \\ &(1 + 2I)^2 = 1 + 4J, \\ &(1 + 4J)^2 = 1 + 8K, \\ &(1 + 8K)^2 = 1 + 16L, \end{aligned}$$

respectivamente.

**14.**
$$\begin{aligned} &x + x^9 + x^{25} + x^{49} + x^{81} + \ldots \\ &= x(1 + x^8 + x^{24} + x^{48} + x^{80} + \ldots) = xP, \end{aligned}$$

donde $P$ denota una serie de valor en que los coeficientes de $x^n$ se desvanecen cuando $n$ no es divisible por 8. Las funciones generatrices de

$$S_1(n), \quad S_2(n), \quad S_4(n), \quad S_8(n)$$

son

$$xP, \quad x^2P^2, \quad x^4P^4, \quad x^8P^8,$$

respectivamente.

**15.** Desde los ejemplos 6 y 11,

$$G^{k+l} = G^kG^l$$

donde $G$ vale para la función generativa de $R_1(n)$.

**16.** Análogo al ejemplo 15, desde los ejemplos 6 y 12.

**17.** Usar los ejemplos 15 y 16 con $k = l = 4$. El cálculo real fue realizado por este método, con comprobaciones ocasionales por otros métodos, como el ejemplo 4.16 y el 23.

**18.** (1) De los ejemplos 12 y 16 se sigue

$$S_4(4)S_4(8n-4) + S_4(12)S_4(8n-12) + \ldots + S_4(8n-4)S_4(4) = S_8(8n)$$

Se conjeturó en la sección 4.6 que $S_4(4(2n-1)) = \sigma(2n-1)$ y en el ejemplo 4.23 que $S_8(8u) = \sigma_3(u)$ si $u$ es un número impar.

(2) $$\begin{aligned}\sigma(1)\sigma(9) + \sigma(3)\sigma(7) + \sigma(5)\sigma(5) + \sigma(7)\sigma(3) + \sigma(9)\sigma(1) \\ = 2(1 \times 13 + 4 \times 8) + 6 \times 6 \\ = 126 = 5^3 + 1^3 = \sigma_3(5).\end{aligned}$$

(3) Parece razonable que tal verificación aumenta nuestra confianza en ambas conjeturas, en algún grado.

**19.** $$\sum \left[\frac{u-1}{2} - 5\,\frac{k(k+1)}{2}\right] s_{u-k(k+1)} = 0$$

para $u = 1, 3, 5, \ldots$; la suma se extiende sobre todos los enteros no-negativos $k$ que satisfacen la desigualdad

$$0 \leqq k(k+1) < u.$$

**20.** $$\begin{aligned}\sigma(3) &= 4\sigma(1) \\ 2\sigma(5) &= 3\sigma(3) \\ 3\sigma(7) &= 2\sigma(5) + 12\sigma(1) \\ 4\sigma(9) &= \sigma(7) + 11\sigma(3).\end{aligned}$$

La última es cierta, puesto que

$$4 \times 13 = 8 + 11 \times 4.$$

**21.** La fórmula periódica ha sido probada para $S_4(4(2n-1))$ en el ejemplo 19. Esta fórmula periódica significa, en efecto, un sistema infinito de relaciones que determina $S_4(4(2n-1))$ si $S_4(4)$ es dado. Ahora, nosotros sabemos que

$$S_4(4) = \sigma(1) = 1.$$

Si $\sigma(2n-1)$ satisface el mismo sistema de relaciones como $S_4(4(2n-1))$,

$$S_4(4(2n-1)) = \sigma(2n-1)$$

para $n = 1, 2, 3, \ldots$, porque el sistema no es ambiguo. Si, recíprocamente, la última ecuación vale, $\sigma(2n - 1)$ satisface dichas relaciones recursivas.

**22.** Asumimos que

$$G = a_0 + a_1x + a_2x^2 + a_3x^3 + \ldots,$$
$$H = u_0 + u_1x + u_2x^2 + u_3x^3 + \ldots,$$
$$G^k = H.$$

Se sigue, como en el ejemplo 19, que

$$GxH' - kxG'H = 0.$$

Igualar a 0 el coeficiente de $x^n$:

$$\sum_{m=0}^{n} [n - (k + 1)m]a_m u_{n-m} = 0.$$

Considerar $a_0, a_1, a_2, \ldots$ como dados. A partir de la última ecuación podemos expresar $u_n$ en términos de $u_{n-1}, u_{n-2}, \ldots u_1, u_0$ teniendo en cuenta que $a_0 \neq 0$. Observar que $u_0 = a_0^k$.

**23.** Aplicar el ejemplo 22 al caso $k = 8$,

$$G = 1 + 2x + 2x^4 + 2x^9 + 2x^{16} + \ldots$$

Por el ejemplo 11 el resultado del ejemplo 22 da

$$nR_8(n) = 2(9 - n)R_8(n - 1) + 2(36 - n)R_8(n - 4) + 2(81 - n)R_8(n - 9) + \ldots$$

Tenemos $R_8(n)/16 = r_n$. Luego, $r_0 = 1/16$ y encontramos $r_1, r_2, r_3, \ldots r_{10}$, sucesivamente,

$$\begin{aligned}
r_1 &= 16r_0 \\
2r_2 &= 14r_1 \\
3r_3 &= 12r_2 \\
4r_4 &= 10r_3 + 64r_0 \\
5r_5 &= 8r_4 + 62r_1 \\
6r_6 &= 6r_5 + 60r_2 \\
7r_7 &= 4r_6 + 58r_3 \\
8r_8 &= 2r_7 + 56r_4 \\
9r_9 &= 54r_5 + 144r_0 \\
10r_{10} &= -2r_9 + 52r_6 + 142r_1.
\end{aligned}$$

Al utilizar numéricamente estas fórmulas tenemos que comprobar algo importante: el lado derecho de la ecuación que da $r_n$ debe ser divisible por $n$.

El mismo método da una fórmula de recursión para $R_k(n)$, $k$ siendo cualquier entero dado $\geqq 2$, y también para $S_k(n)$.

**25.** Llamamos $s$ el producto infinito. Calculando $-xd \log s/dx$ y utilizando el número 10 de la memoria de Euler, obtenemos

$$k\,\Sigma\sigma(l)x^l = \frac{\Sigma n a_n x^n}{1 - \Sigma a_m x^m};$$

los límites para las tres sumas son 1 e $\infty$. Multiplicar con el denominador del lado derecho y enfocar el coeficiente de $x^n$.

El caso de Euler es $k = 1$. También el caso $k = 3$ da un resultado relativamente simple (ver la obra de HARDY y WRIGHT, citada en el ejemplo 24, págs. 282 y 283, teoremas 355 y 357). En otros casos no sabemos bastante sobre la ley $a_n$.

Sin solución: **1, 24.**

## CAPITULO VII

**1.** [Standford, 1950]

$$1 - 4 + 9 - 16 + \ldots + (-1)^{n-1} n^2 = (-1)^{n-1} \frac{n(n+1)}{2}.$$

El paso de $n$ a $n + 1$ se requiere para verificar que

$$(-1)^n (n+1)^2 = (-1)^n \frac{(n+1)(n+2)}{2} - (-1)^{n-1} \frac{n(n+1)}{2}.$$

**2.** Para probar

$$P_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2},$$

$$S_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3},$$

$$S_n^{\infty} = 2\binom{n}{0} + 2\binom{n}{2},$$

utilizamos los ejemplos 3.11 y 3.12 combinados con la expresión de $P_n$, y el ejemplo 3.20, respectivamente. Luego, dando por supuestas las expresiones anteriores, se nos pide verificar

$$P_{n+1} - P_n = \binom{n}{0} + \binom{n}{1}$$

$$S_{n+1} - S_n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2}$$

$$S^{\infty}_{n+1} - S^{\infty}_n = 2\binom{n}{1}.$$

Estas tres se siguen del conocido hecho (la relación básica en el triángulo de Pascal) de que

$$\binom{n+1}{k+1} - \binom{n}{k+1} = \binom{n}{k}.$$

**3.** *How to Solve It*, págs. 103-110.

**4.** $$\frac{3}{4}, \frac{2}{3} = \frac{4}{6}, \frac{5}{8}, \frac{3}{5} = \frac{6}{10}, \cdots \frac{n+1}{2n}.$$

El paso de $n$ a $n+1$ requiere que verifiquemos que

$$1 - \frac{1}{(n+1)^2} = \frac{n+2}{2n+2}\,\frac{2n}{n+1}.$$

Cf. el ejemplo 2.23.

**5.** $$-\frac{3}{1}, -\frac{5}{3}, -\frac{7}{5}, -\frac{9}{7}, \cdots -\frac{2n+1}{2n-1}.$$

El paso de $n$ a $n+1$ requiere que verifiquemos que

$$1 - \frac{4}{(2n+1)^2} = \frac{2n+3}{2n+1}\,\frac{2n-1}{2n+1}.$$

Cf. el ejemplo 2.31.

**6.** El caso general es, en efecto, equivalente al caso límite

$$\frac{x}{1-x} = \frac{x}{1+x} + \frac{2x^2}{1+x^2} + \frac{4x^4}{1+x^4} + \frac{8x^8}{1+x^8} + \cdots$$

a partir de cuyo caso particular propuesto podemos derivar así: sustituir $c^{16}$ para $x$ y multiplicar por 16, obteniendo

$$\frac{16x^{16}}{1-x^{16}} = \frac{16x^{16}}{1+x^{16}} + \frac{32x^{32}}{1+x^{32}} + \ldots;$$

luego sustraemos de la ecuación original. Si tenemos $2^{n+1} = m$, el paso de $n$ a $n+1$ requiere

$$\frac{mx^m}{1+x^m} = -\frac{2mx^{2m}}{1-x^{2m}} + \frac{mx^m}{1-x^m}.$$

**7.** Para probar

$$1 + 3 + 5 + 7 + \ldots + 2n - 1 = n^2$$

El paso de $n$ a $n+1$ requiere verificar que

$$2n + 1 = (n+1)^2 - n^2.$$

**8.** El $n$-avo término en la cuarta fila de la tabla es

$$\begin{aligned}(1+2) + (4+5) + \ldots + (3n-5+3n-4) + 3n - 2 \\ = 3 + 9 + \ldots + [6(n-1) - 3] + 3n - 2 \\ = 6\frac{n(n-1)}{2} - 3(n-1) + 3n - 2 = 3n^2 - 3n + 1.\end{aligned}$$

El paso de $n-1$ a $n$ requiere, en efecto,

$$n^3 - (n-1)^3 = 3n^2 - 3n + 1.$$

**9.** Tras $n^2$, $n^3$ y $n^4$, la generalización referente a $n^k$ es obvia. El sencillo caso de $n^2$ era conocido desde la Antigüedad; Alfred Moessner descubrió el resto bastante recientemente por inducción empírica, y Oskar Perron lo probó por inducción matemática. Ver *Sitzungsberichte der Bayerischen Akademie der Wissenschaften*, Math.-naturwissenschaftliche Klasse, 1951, págs. 29-43.

**10.** Para $k = 1$ el teorema se reduce a la identidad obvia

$$1 - u = -(n-1).$$

El paso de $k$ a $k+1$ requiere que verifiquemos que

$$(-1)^{k+1}\binom{n}{k+1} = (-1)^{k+1}\binom{n-1}{k+1} - (-1)^k\binom{n-1}{k}$$

que es la relación básica del triángulo de Pascal, ya encontrada en el ejemplo 2.

**11.** [Standford, 1946]. Llamemos $P_n$. al número pedido de emparejamientos de $2n$ jugadores. El $n$-avo jugador puede ser emparejado con cualquiera de los otros $2n - 1$ jugadores. Una vez elegido su antagonista quedan

$$2n - 2 = 2(n - 1)$$

jugadores que pueden ser emparejados de $P_{n-1}$ maneras. De aquí que

$$P_n = (2n - 1)P_{n-1}.$$

**12.** Llamemos $A_n$ al enunciado propuesto para ser probado, referido a $f_n(x)$. En vez de $A_n$ probaremos $A'_n$.

$A'_n$. La función $f_n(x)$ es un cociente *cuyo denominador es* $(1 - x)^{n+1}$ y el numerador un polinomio de *grado n*, cuyo término constante es 0, y el otro coeficiente enteros positivos.

Obsérvese que $A'_n$ afirma más que $A_n$; los puntos en que $A'_n$ va más allá de $A_n$ están señalados por medio de itálicas. Dando por supuesto $A'_n$, establecemos

$$(1 - x)^{n+1} f_n(x) = P_n(x) = a_1 x + a_2 x^2 + \ldots + a_n x^n,$$

donde se supone que $a_1$, $a_2$, ... *an* son enteros positivos. De la definición periódica derivamos la fórmula

$$P_{n+1}(x) = x[(1 - x)P'_n(x) + (n + 1)P_n(x)]$$

y esto muestra que los coeficientes de $x$, $x^2$, $x^3$, ... $x^n$ y $x^{n+1}$ en $P_{n+1}(x)$ son $a_1$, $na_1 + 2a_2$, $(n - 1)\ a_2 + 3a_2$, ... $2a_{n+1} + na_n$, $a_n$, respectivamente, lo cual hace obvia la aserción.

**13.** (1) La suma de todos los coeficientes de $P_n(x)$ es $n!$. En efecto, esta suma es $P_n(1)$ y la fórmula periodica arroja

$$P_{n+1}(1) = (n + 1)P_n(1).$$

(2) $P_n(x)/x$ es un polinomio recíproco o

$$P_n(1/x)x^{n+1} = P_n(x).$$

En efecto, suponer que $a_1 = a_n$, $a_2 = a_{n-1}$ ...; las correspondientes relaciones para los coeficientes de $P_{n+1}(x)$ se siguen de su expresión dada al final de la solución del ejemplo 12.

**16.** $Q_1 = 1, \quad Q_2 = 3, \quad Q_3 = 45, \quad Q_4 = 4725$

$Q_2/Q_3 = 3, \quad Q_1/Q_2 = 15, \quad Q_3/Q_4 = 105$

sugiere

$$Q_n = 1^n 3^{n-1} 5^{n-2} \ldots (2n-3)^2 (2n-1)^1.$$

En efecto, la definición da

$$\frac{Q_{n+1}}{Q_{n-1}} = \frac{Q_n Q_{n+1}}{Q_{n-1} Q_n} = \frac{(2n)!\,(2n+1)!}{(n!\,2^n)^2} = [1 \cdot 3 \cdot 5 \ldots (2n-1)]^2 (2n+1)$$

y de aquí probamos la ley general por inferencia de $n - 1$ a $n + 1$. Observar que

$$\frac{2n!}{n!2^n} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \ldots (2n-1)2n}{2 \quad . \quad 4 \quad . \quad 6 \ldots \quad . \quad 2n}.$$

**17.** El razonamiento que nos conduce de 3 a 4 se aplica al paso de $n$ a $n + 1$ con una excepción: se rompe en el paso de 1 a 2.

**18.** De $n = 3$ a $n + 1 = 4$: considerar las líneas *a*, *b*, *c* y *d*. Considerar primero el caso en que hay dos líneas diferentes entre las cuatro dichas, por ejemplo, *b* y *c*. Luego el punto de intersección de *b* y *c* está únicamente determinado y debe también valer para *a* (puesto que el enunciado se mantiene para $n = 3$), y también para *d* (por la misma razón). Por tanto, el enunciado se mantiene para $n + 1 = 4$. Si, no obstante, no hay dos diferentes entre las cuatro líneas, el enunciado es obvio. Este razonamiento rompe el paso de 2 a 3.

Sin solución: **14, 15.**

## CAPITULO VIII

**1.** (1) Línea recta, (2) perpendicular, (3) perpendicular común, (4) segmentos de línea a través del punto y el centro dado (EUCLIDES, III, 7, 8), (5) segmento de perpendicular a través del centro, ninguno máximo, (6) segmentos de línea que une los centros. El caso en que la distancia mínima es 0 ha sido descartado.

**2.** (1) Línea recta, (2) perpendicular, (3) perpendicular común, (4) perpendicular, (5) perpendicular común, (6) ver la sección 4, (7) segmentos de la línea que une el punto al centro, (8) segmento de la perpendicular a través del centro, que no sea máximo, (9) segmento de la

perpendicular a través del centro, que no sea máximo, (10) segmentos de la línea que une los centros. El caso en que la distancia mínima es 0 ha sido descartado.

**3.** (1) círculos concéntricos, (2) líneas paralelas, (3) círculos concéntricos.

**4.** (1) esferas concéntricas, (2) planos paralelos, (3) cilindros coaxiales, (4) esferas concéntricas.

**5.** (2) Ver la sección 3. Otras semejantes.

**6.** (6) Hay sólo un cilindro, que tiene la primera línea dada como eje y la segunda como tangente. El punto de contacto es el punto final de la distancia más corta. Otras semejantes.

**7.** Llamemos base a uno de los lados dados. Mantener la base en posición fija, giremos el otro lado sobre el punto final y llamemos a su otro punto final $X$. La trayectoria prescrita de $X$ es un círculo, las líneas de nivel son paralelas a la base, el triángulo con área máxima es un triángulo rectángulo (lo cual es obvio).

**8.** Llamemos base al lado dado, mantengamos ésta en posición fija, llamemos el vértice opuesto $X$ y movámosle. La trayectoria prescrita de $X$ es una elipse, las líneas de nivel son paralelas a la base y el triángulo con área máxima es isósceles.

**9.** Las líneas de nivel son líneas rectas $x + y = \text{const}$, la trayectoria prescrita es (una rama de) una hipérbole de ecuación $xy = A$, donde $A$ es el área dada. Por simetría vemos claramente que hay un punto de contacto con $x = y$.

**10.** Considerar círculos concéntricos extendiéndose con el punto dado como centro. Parece por intuición que hay un primer círculo que toca la línea dada; su radio es la distancia más corta. Esto es cierto en los casos del ejemplo 1 (2) y (4)

**11.** Cruzando puntos medios pasamos de un lado de la línea de nivel a otro y sobre un lado $f$ se toman los valores más altos; mientras sobre el otro, los más bajos con respecto al punto de cruce.

**12.** Usted puede, pero no necesita. El punto más alto puede ser el pico $P$ (puede usted desear ver la vista), o el paso $S$ (puede usted cruzarlo andando de un lado a otro), o el punto inicial de su trayectoria, o el punto final, o un punto angular de ésta.

**13.** (1) La línea de nivel para 180° es el segmento $AB$, la línea de nivel para 0° es la línea recta que pasa por $A$ y $B$, excepto el seg-

mento $AB$. Cualquier otra línea de nivel consta de dos arcos circulares, ambos con puntos finales $A$ y $B$, y simétricos uno a otro con recpecto a la línea recta que pasa por $A$ y $B$. (2) Si dos líneas de nivel son diferentes, una está dentro de la otra; $< AXB$ alcanza un valor más alto sobre la línea de nivel interna y uno más bajo sobre la externa. Mediante una interpretación adecuada esto se aplica también a 0°.

**14.** El mínimo se alcanza en el punto en que la línea $l$ cruza la línea que pasa por $A$ y $B$, es decir, una línea de nivel. Esto no contradice el principio del ejemplo 11; a ambos lados de esta particular línea de nivel $< AXB$ toma valores más altos que sobre ella.

**15.** La notación como en el ejemplo 6. Hagamos a $c$ constante por un momento. Luego, puesto que $V = abc/3$ es dado, también $ab$ es constante, y

$$S = 2ab + 2(a + b)c$$

es un mínimo, cuando 2 ($a$ + b), el perímetro de un rectángulo con área dada $ab$, es un mínimo. Esto ocurre cuando $a = b$. Ahora, cambiemos el punto de vista y mantengamos otra arista constante.

**16.** Mantengamos uno de estos lados constantes. Luego tenemos el caso del ejemplo 8, y los otros dos lados deben ser iguales (el triángulo es isósceles). Dos lados cualesquiera son accesibles a este argumento, y el triángulo deberá ser equilátero.

**17.** Mantengamos fijos el plano de la base y el vértice opuesto, pero movamos la base que es un triángulo inscrito en un círculo dado. La altura es constante; el área de la base (y con ella el volumen) alcanzan un máximo cuando la base es equilátera, por la sección 4 (2). Podemos elegir cualquier lado como base, y cada lado debe ser equilátero, y, en consecuencia, el tetraedro regular cuando se alcanza el volumen máximo.

**18.** Considerar el triángulo entre $a$ y $b$ como la base. Sin cambiar la altura correspondiente, cambiar la base en un triángulo rectángulo; este cambio aumenta el área de la base (ejemplo 7) y, en consecuencia, el volumen. Podríamos tratar ahora otro par de lados de modo semejante.

**19.** Fijando el punto final sobre el cilindro, encontramos, por el ejemplo 2 (7), que la distancia más corta es perpendicular a la esfera. Fijando el punto final sobre la esfera podemos convencernos de que la distancia más corta es también perpendicular al cilindro. Por tanto, debe ser perpendicular a ambos. Esto puede mostrarse también de modo directo.

**20.** El procedimiento del ejemplo 19 muestra que la distancia más corta debe ser perpendicular a ambos cilindros. En efecto, ésta cae en la misma línea como la distancia más corta entre los ejes de los cilindros; ver la sección 4 (1).

**21.** El procedimiento del ejemplo 19 y un caso análogo del ejemplo 10 en el espacio.

**22.** Por hipótesis,

$$f(X, Y, Z, \ldots) \leqq f(A, B, C, \ldots)$$

para todos los valores admisibles de $X$, $Y$, $Z$, . . . Por tanto, en particular,

$$f(X, B, C, \ldots) \leqq f(A, B, C, \ldots)$$
$$f(X, Y, C, \ldots) \leqq f(A, B, C, \ldots),$$

y así sucesivamente; $X$, $Y$, $Z$, ... pueden ser variables números, longitudes, ángulos, puntos, ...

**24.** O $x_1 = y_1 = z_1$ (caso excepcional), o, para $n \geqq 1$, de los tres valores $x_n$, $y_n$, $z_n$ sólo dos son diferentes. Llamemos $d_n$ el valor absoluto de la diferencia; por ejemplo,

$$d_1 = |x_1 - z_1|, \qquad d_2 = |y_2 - x_2|.$$

Por definición,

$$\pm d_2 = x_2 - y_2 = \frac{z_1 + x_1}{2} - y_1 = \frac{z_1 + x_1}{2} - z_1$$

$$= \frac{x_1 - z_1}{2} = \pm \frac{d_1}{2}.$$

o $d_2 = d_1/2$. Del mismo modo,

$$d_n = d_{n-1}/2 = d_{n-2}/4 = \ldots = d_1/2^{n-1}$$

y así,

$$\left| x^n - \frac{l}{3} \right| = \left| x_n - \frac{n_n + y_n + z_n}{3} \right| = \frac{|x_n - y_n + x^u - z^n|}{3}$$

$$\leqq 2d_n/3 \to 0.$$

**25.** Consideremos $n$ números positivos $x_1, x_2, \ldots x_n$ con la media aritmética dada $A$,

$$x_1 + x_2 + \ldots + x_n = nA.$$

Si $x_1, x_2, \ldots x_n$ no son todos iguales, uno de ellos, digamos $x_1$, es el más pequeño, y otro, digamos $x_2$, es el más grande. (La elección de los suscritos es una inicua simplificación, sólo una cuestión de notación adecuada. *No* necesitamos hacer la injustificable asunción de que *sólo* $x_1$ toma el valor más pequeño.) Luego,

$$x_1 < A < x_2.$$

Hagamos ahora

$$x_1' = A, \qquad x_2' = x_1 + x_2 - A, \qquad x_3' = x_3, \ldots x_n' = x_n.$$

Luego,

$$x_1 + x_2 + \ldots + x_n = x_1' + x_2' + \ldots + x_n'$$

y

$$x_1'x_2' - x_1x_2 = Ax_1 + Ax_2 - A^2 - x_1x_2 = (A - x_1)(x_2 - A) > 0.$$

Por tanto,

$$x_1x_2x_3 \ldots x_n < x_1' x_2'x_3' \ldots x_n'.$$

Si $x_1', x_2', \ldots x_n'$ no son todos iguales, repetimos el proceso obteniendo otro conjunto de $n$ números $x_1'', x_2'', \ldots x_n''$ tal que

$$x_1' + x_2' + \ldots + x_n' = x_1'' + x_2'' + \ldots + x_n''$$
$$x_1'x_2'x_3' \ldots x_n' < x_1''x_2''x_3'' \ldots x_n''.$$

El conjunto $x_1', x_2', \ldots x_n'$ tiene al menos un término igual a $A$, el conjunto $x_1'', x_2'', \ldots x_n''$ tiene al menos dos términos iguales a $A$. Por lo menos, el conjunto $x_1^{(n-1)}, x_2^{(n-1)}, \ldots x_n^{(n-1)}$ contendrá $n - 1$ y, por tanto, $n$ términos iguales a $A$, y así,

$$x_1x_2 \ldots x_n < x_1^{(n-1)}x_2^{(n-1)} \ldots x_n^{(n-1)}$$
$$= A^n = \left(\frac{x_1 + x_2 + \ldots + x_n}{n}\right)^n.$$

**26.** Conectando el punto inicial común de las perpendiculares $x$, $y$, $z$ con los tres vértices del triángulo, subdividimos el último dentro de los tres triángulos más pequeños. Expresando que la suma de las áreas de estas tres partes es igual al área del conjunto, obtenemos

$x + y + z = l$. La ecuación $x = \text{const}$ se representa por una línea paralela a la base del triángulo equilátero, la ecuación $y = z$ por la altura. El primer segmento de la línea quebrada en la figura 8.9 es paralelo a la base y termina en la altura. El primer paso del ejemplo 25 es representado por un segmento paralelo a la base y termina en la línea con ecuación $y = l/3$, que es paralelo a otro lado y pasa por el centro. El segundo paso es representado por un segmento a lo largo de la línea $y = l/3$ y termina en el centro.

**27.** Para la solución modelada tras el ejemplo 25 ver RADEMACHER-TOEPLITZ, págs. 11-14, 114-117.

**28.** Variación parcial.

**29.** En el punto donde se alcanza el extremo las ecuaciones

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = 0, \qquad \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = 0$$

se mantienen con un valor adecuado de $\lambda$. Esta condición se deriva bajo la asunción de que $\partial g/\partial x$ y $\partial g/\partial y$ no desaparecen. Bajo la nueva asunción de que $\partial f/\partial x$ y $\partial f/\partial y$ no desaparecen, las ecuaciones con $\lambda$ expresan que la curva $g = 0$ (la trayectoria prescrita) es tangente a la curva $f = \text{const}$ (una línea de nivel) que pasa por el punto del extremo en este punto.

**30.** En un pico, o en un paso, $\partial f/\partial x = \partial f/\partial y = 0$. En un punto angular de la trayectoria prescrita, $\partial g/\partial x$ y $\partial g/\partial y$ (si existen) son $= 0$. Un extremo en el punto inicial (o final) de la trayectoria prescrita no cae en absoluto bajo la condición analítica citada en el ejemplo 29, que se refiere a un extremo relativo a *todos* los puntos $(x, y)$ en cierta proximidad, satisfaciendo $g(x, y) = 0$.

**31.** La condición es

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} = \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} = \frac{\partial f}{\partial z} + \lambda \frac{\partial g}{\partial z} = 0.$$

Se asume que las tres derivadas parciales de $g$ no son 0. Bajo la nueva asunción de que las tres derivadas parciales de $f$ no son 0, las tres ecuaciones expresan que la superficie $g = 0$ y $f = \text{const}$ que pasan por el punto del extremo son tangente entre sí en ese punto.

**32.** La condición consta de tres ecuaciones, de las cuales la primera, relativa al eje de las $x$, es

$$\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} + \mu \frac{\partial h}{\partial x} = 0.$$

Se asume que tres determinantes, de la cuales la primera es

$$\frac{\partial g}{\partial y}\frac{\partial h}{\partial z} - \frac{\partial g}{\partial z}\frac{\partial h}{\partial y}$$

no desaparecen. Bajo la nueva asunción de que las tres derivadas parciales de $f$ no son 0, las tres ecuaciones expresan que la curva cuya intersección de las dos superficies $g = 0$ y $h = 0$, es tangente a la superficie $f = \text{const}$ pasa por el punto del extremo en ese punto.

**34.** La conclusión deseada es: el cubo sólo alcanza el máximo. Por tanto, cuando la desigualdad alcanza una igualdad, el cubo debe aparecer; es decir, tendríamos $x = y$, o (concerniente a las áreas) $x^2 = xy$. Sin embargo, la desigualdad utilizada (sin éxito) alcanza una igualdad cuando $2x^2 = 4xy$: podíamos haber predicho desde ahora que fracasará.

Mirando, o sin mirar, la sección 6, dividimos $S$ en tres pares de caras opuestas

$$S = 2x^2 + 2xy + 2xy$$

y aplicar el teorema de las medias:

$$(S/3)^3 \geqq 2x^2 \cdot 2xy \cdot 2xy = 8x^4y^2 = 8V^2.$$

La igualdad se alzanza si, y sólo si, $2x^2 = 2xy$, o $x = y$; es decir, sólo por el cubo.

Moraleja: prever el caso de igualdad puede servir de guía para decidirnos puede darnos una sugenrecia.

**35.** $V$, $S$, $x$ e $y$ denotan el volumen, la superficie, el radio y la altura del cilindro, respectivamente, de modo que

$$V = \pi x^2 y, \qquad S = 2\pi x^2 + 2\pi xy.$$

La conclusión deseada, $y = 2x$, guía nuestra elección: con

$$S = 2\pi x^2 + \pi xy + \pi xy$$

el teorema de las medias da

$$(S/3)^3 \geqq 2\pi x^2 \cdot \pi xy \cdot \pi xy = 2\pi^3 x^4 y^2 = 2\pi V^2,$$

con igualdad sólo para $y = 2x$.

**36.** El ejemplo 34 es un caso particular, el ejemplo 35 un caso límite. $V$, $S$, $y$ y $x$ denotan el volumen, la superficie, la altura del prisma y la longitud de un cierto lado de su base, respectivamente. $a$ y $l$ denotan el área y el perímetro, respectivamente, de un polígono similar a la base cuyo lado correspondiente al lado de longitud $x$ de la base es de longitud 1. Luego,

$$V = ax^2y, \qquad S = 2ax^2 + lxy.$$

En los ejemplos 34 y 35 se obtiene el máximo de $S$, cuando el área de la base (ahora $ax^2$) es $S/6$. Esperando que esto valga también para el presente caso general, tenemos una cuestión; pongamos

$$S = 2ax^2 + lxy/2 + lxy/2,$$

y obtenemos, utilizando el teorema de la medias,

$$(S/3)^3 \geqq 2ax^2 \cdot (lxy)^2/4 = [l^2/(2a)]\, V^2$$

con igualdad si $ax^2 = lxy/4 = S/6$.

**37.** $V$. $S$, $x$ e $y$ denotan el volumen de la pirámide doble, su superficie, un lado de su base y la altura de una de las pirámides constituyentes, respectivamente. Luego,

$$V = 2x^2\, y/3, \qquad S = 8x\, [(x/2)^2 + y^2]^{1/2}/2.$$

En el caso del octaedro regular la doble altura de una pirámide constituyente es igual al diámetro de la base, o

$$2y = 2^{1/2}x, \quad \text{o} \quad 2y^2 = x^2$$

Habiendo obtenido esta clave ponemos

$$S^2 = 4x^2(x^2 + 2y^2 + 2y^2),$$

$$(S^2/3)^3 \geqq 4^3x^6x^22y^22y^2 = 4^4x^8y^4\, (6V)^4.$$

La igualdad se da sólo si $x^2 = 2y^2$. Notemos que, en este caso, $S = 3^{1/2}\, 2x^2$.

**38.** $V$, $S$, $x$ e $y$ denotan el volumen del doble cono, su superficie, el radio de su base y la altura de uno de los conos constituyentes, respectivamente. Luego,

$$V = 2\pi x^2y/3, \qquad S = 2 \cdot 2\pi x(x^2 + y^2)^{1/2}/2.$$

Considerar el triángulo rectángulo de lados $x$, $y$ y la hipotenusa $(x^2 + y^2)^{1/2}$. Si la proyección de $x$ sobre la hipotenusa es 1/3 del último (como esperamos que sea en el caso mínimo),

$$x^2 = (x^2 + y^2)/3$$

o $2x^2 = y^2$. Habiendo obtenido esta sugestión tenemos

$$S^2 = 2\pi^2 x^2(2x^2 + y^2 + y^2),$$
$$(S^2/3)^3 \geqq 8\pi^6 x^6 \cdot 2x^2y^2y^2 = \pi^2(3V)^4.$$

La igualdad sólo se da si $2x^2 = y^2$. Notemos que, en este caso, $S = 3^{1/2}\, 2 \cdot \pi x^2$.

**39.** $V$, $S$ e $y$ denotan el volumen de la doble pirámide, su superficie y la altura de una de las pirámides constituyentes, respectivamente. Conectemos $x$, $a$ y $l$ con la base de la doble pirámide del mismo modo que en la solución del ejemplo 36 con la base del prisma. $p$ denota el radio del círculo inscrito en la base. Luego,

$$V = 2ax^2y/3,$$
$$ax^2 = lxp/2,$$
$$S = 2lx(p^2 + y^2)^{1/2}/2 = (4a^2x^4 + l^2x^2y^2)^{1/2}.$$

En los ejemplos 37 y 38, $S$ es un mínimo cuando $S = 3^{1/2}2ax^2$, que da

$$l^2x^2y^2 = 8a^2x^4.$$

Por esta sugerencia ponemos

$$S^2 = 4a^2x^4 + l^2x^2y^2/2 + l^2x^2y^2/2,$$
$$(S^2/3)^2 \geqq 4a^2x^4 \cdot (l^2x^2y^2)^2/4 = (l^4/a^2)\,(3V/2)^4.$$

La igualdad sólo se da si, y sólo si, la base $ax^2 = S/(3^{1/2}2)$.

**40.** Hay una conjetura plausible: el triángulo equilátero tiene el perímetro mínimo para un área dada, o el área máxima para un perímetro dado. $a$, $b$, $c$, $A$ y $L = 2p$ denominan los tres lados del triángulo, su área y la longitud de su perímetro, respectivamente. Por la fórmula de Herón,

$$A^2 = p(p - a)\,(p - b)\,(p - c).$$

El empleo del teorema de las medias es fuertemente sugerido: $A$ no será demasiado largo, cuando $p$ es dado; el lado derecho de la ecuación es un producto. ¿Cómo aplicaremos el teorema? He aquí una indica-

ción: *si* el triángulo es equilátero, $a = b = c$, o $p - a = p - b = p - c$. Por tanto, procedemos como sigue:

$$A^2/p = (p-a)(p-b)(p-c)$$
$$\leqq \left(\frac{p-a+p-b+p-c}{3}\right)^3$$
$$= (p/3)^3.$$

Es decir, $A^2 \leqq L^4/(2^4 3^3)$, y hay igualdad sólo en el caso del triángulo equilátero. Cf. el ejemplo 16.

**41.** Hay una conjetura plausible: el cuadrado.

$a$ y $b$ comprenden el ángulo $\Phi$, $c$ y $d$ el ángulo $\Psi$, y $\Phi + \Psi = \varepsilon$. Obtenemos

$$2A = ab \operatorname{sen} \Phi + cd \operatorname{sen} \psi.$$

Expresando de dos maneras diferentes la diagonal del cuadrado que separa $\Phi$ y $\psi$, obtenemos

$$a^2 + b^2 - 2ab \cos \Phi = c^2 + d^2 - 2cd \cos \Psi.$$

Tenemos ahora tres relaciones para eliminar $\Phi$ y $\Psi$. Sumando

$$(a^2 + b^2 - c^2 - d^2)^2 = 4a^2b^2 \cos^2 \Phi + 4c^2d^2 \cos^2 \Psi - 8abcd \cos \Phi \cos \Psi$$
$$16A^2 = 4a^2b^2 \operatorname{sen}^2 \Phi + 4c^2d^2 \operatorname{sen}^2 \Psi + 8abcd \operatorname{sen} \Phi \operatorname{sen} \Psi$$

obtenemos

$$16A^2 + (a^2 + b^2 - c^2 - d^2)^2 = 4a^2b^2 + 4c^2d^2 - 8abcd \cos \varepsilon$$
$$= 4(ab + cd)^2 - 16abcd (\cos \varepsilon/2)^2.$$

Finalmente, observando las diferencias de los cuadrados y poniendo

$$a + b + c + d = 2p = L,$$

encontramos

$$A^2 = (p-a)(p-b)(p-c)(p-d) - abcd (\cos \varepsilon/2)^2.$$

En el caso probable de igualdad (el cuadrado) los lados son iguales,

y también lo son las cantidades $p-a, p-b, p-c, p-d$. Con esta indicación obtenemos

$$\begin{aligned} A^2 &\leqq (p-a)(p-b)(p-c)(p-d) \\ &\leqq \left(\frac{p-a+p-b+p-c+p-d}{4}\right)^4 \\ &= (p/2)^4 = (L/4)^4. \end{aligned}$$

Al objeto de que ambas desigualdades encontradas lleguen a ser igualdades debemos hacer $\varepsilon = 180°$ y $a = b = c = d$.

**42.** El prisma es mucho más accesible que los otros dos sólidos que tantearemos, tras cuidadosa preparación, en los ejemplos 46 y 47. $L$ denomina el perímetro de la base y $h$ la altura del prisma. Cualquier cara lateral es un paralelogramo; su base es un lado de la base del sólido, y su altura $\geqq h$. Por tanto, la superficie lateral del prisma es $\geqq Lh$, y la igualdad se obtiene si, y sólo si, todas las caras laterales son perpendiculares a la base y por ello el prisma es un prisma recto.

**43.** $x_j$, $y_j$ sean las cordenadas de $P_j$, para $j = 0, 1, 2, \ldots n$, y pongamos

$$x_j = x_{j-1} + u_j, \qquad y_j = y_{j-1} + v_j$$

para $j = 1, 2, \ldots n$. Luego el lado izquierdo de la deseada desigualdad es la longitud de la línea quebrada $P_0P_1P_2 \ldots P_n$ y el lado derecho de la longitud de la línea recta $P_0P_n$, que es la distancia más corta entre sus puntos finales.

**44.** En el caso $n = 2$ examinamos (la notación está ligeramente cambiada) la aserción

$$(u^2 + v^2)^{1/2} + (U^2 + V^2)^{1/2} \geqq [(u + U)^2 + (v + V)^2]^{1/2}$$

Transformamos esto en formas equivalentes mediante cuadrados y otras operaciones algebraicas:

$$\begin{aligned} (u^2 + v^2)^{1/2}(U^2 + V^2)^{1/2} &\geqq uU + vV, \\ u^2V^2 + v^2U^2 &\geqq 2uvUV, \\ (uV - vU)^2 &\geqq 0. \end{aligned}$$

En la última forma la aserción es obviamente cierta. La igualdad se alcanza si, y sólo si,

$$u : v = U : V.$$

Nos enfrentamos con el caso $n = 3$ aplicando repetidamente el caso $n = 2$:

$$(u_1^2 + v_1^2)^{1/2} + (u_2^2 + v_2^2)^{1/2} + (u_3^2 + v_3^2)^{1/2}$$

$$\geqq [(u_1 + u_2)^2 + (v_1 + v_2)^2]^{1/2} + (u_3^2 + v_3^2)^{1/2}$$

$$\geqq [(u_1 + u_2 + u_3)^2 + (v_1 + v_2 + v_3)^2]^{1/2}.$$

Y así sucesivamente para $n = 4, 5, \ldots$ En efecto, hemos usado la inducción matemática.

**45.** Sea $h$ la altura y dividamos la base por el pie de la altura en dos segmentos, de longitud $p$ y $q$, respectivamente. Tenemos que probar que

$$(p^2 + h^2)^{1/2} + (q^2 + h^2)^{1/2} \geqq 2\left[\left(\frac{p+q}{2}\right)^2 + h^2\right]^{1/2}$$

$$= [(p + q)^2 + (h + h)^2]^{1/2}$$

que es el caso del ejemplo 43. Para la igualdad debemos tener

$$p : h = q : h,$$

o $p = q$, es decir, un triángulo isósceles.

**46.** Sea $h$ la altura de $P$, $a_1, a_2, \ldots a_n$ los lados de la base de $P$, y $p_1, p_2, \ldots p_n$ las perpendiculares desde el pie de la altura sobre los lados respectivos. $\Sigma$ denota una suma con $j$ ordenaciones de $j = 1$ a $j = n$. Luego,

$$A = \Sigma a_j p_j/2$$

$$S = A + \Sigma a_j(p_j^2 + h^2)^{1/2}/2.$$

Estas expresiones son más sencillas para la pirámide recta $P_0$, puesto que todas las perpendiculares desde el pie de la altura sobre los lados tienen un valor común $p_0$. En consecuencia,

$$A_0 = L_0 p_0/2$$

$$S_0 = A_0 + L_0(p_0^2 + h^2)^{1/2}/2$$

$$= A_0 + (4A_0^2 + h^2 L_0^2)^{1/2}/2;$$

$P$ y $P_0$ tienen la misma altura $3V/A = 3V_0/A_0 = b$. Utilizando el ejemplo 43 y nuestras suposiciones, obtenemos

$$\begin{aligned} 2(S - A) &= \Sigma[(a_j p_j)^2 + (a_j h)^2]^{1/2} \\ &\geqq [(\Sigma a_j p_j)^2 + (\Sigma a_j h)^2]^{1/2} \\ &= [4A^2 + h^2 L^2]^{1/2} \\ &\geqq [4A^2 + h^2 L_0^2]^{1/2} \\ &= 2(S_0 - A). \end{aligned}$$

En conseclencia, $S \geqq S_0$. Para la igualdad deben ser igualadas las desigualdades encontradas y para ello deben satisfacerse dos condiciones. Primero,

$$p_1 : h = p_2 : h = \ldots = P_n : h,$$

esto es, $P$ es una pirámide recta. Segundo, $L = L_0$.

**47.** Demos dos pasos: (1) Cambiar la base $D$ en $D_0$, y las dos pirámides en pirámides rectas, dejando sus alturas invariables. Obtenemos así una pirámide doble $D'$ que no es necesariamente una pirámide recta doble. (Sus dos pirámides constituyentes son pirámides rectas, pero quizá de alturas diferentes.) (2) Cambiar $D'$ en $D_0$. El paso (1) puede sólo disminuir la superficie, por el ejemplo 46. Las alturas de las dos pirámides constituyentes de $D'$, de longitudes $h_1$ y $h_2$, caen dentro de la misma línea recta. $p_0$ denota el radio del círculo inscrito en la base de $D_0$. Luego la superficue de $D'$ es

$$\begin{aligned} S &= [(p_0^2 + h_1^2)^{1/2} + (p_0^2 + h_2^2)^{1/2}] L_0/2 \\ &\geqq 2 \left[ p_0^2 + \left( \frac{h_1 + h_2}{2} \right)^2 \right]^{1/2} \frac{L_0}{2} = S_0 \end{aligned}$$

por el ejemplo 45.

**48.** Dejando todo el tiempo el volumen $V$ invariable damos tres pasos: (1) Transformamos el prisma dado en un prisma recto, dejando la base invariable en forma y tamaño. (2) Transformamos la base en un cuadrado, dejando su área $A$ invariable. (3) Transformamos el prisma recto con base cuadrada en un cubo. Los pasos (1) y (3) pueden disminuir la superficie $S$, por los ejemplos 42 y 34, respectivamente. El paso (2) deja la altura $h = V/A$ invariable, y puede sólo disminuir $L$, el perímetro de la base, por el ejemplo 41; pero $S = 2A + Lh$. A menos que

el prisma sea un cubo desde el principio, por lo menos uno de los tres pasos disminuye realmente $S$. El debilitador teorema 34 sirve como piedra de paso.

**49.** Se sigue de los ejemplos 47, 41 y 37 como el anterior ejemplo 48 se sigue a los ejemplos 42, 41 y 34. Sin embargo, podemos combinar con ventaja los dos pasos correspondientes a (1) y (2) del ejemplo 48 en uno solo, gracias a la aguda formulación del ejemplo 47.

**50.** Partamos de una pirámide con base triangular (un tetraedro, no necesariamente regular). Lo transformamos en una pirámide recta, dejando invariable el volumen $V$ y el área de la base $A$, pero cambiando (si es necesario) la base en un triángulo equilátero. Este disminuye el perímetro de la base $L$ por el ejemplo 40 y, en consecuencia, la superficie $S$ por el ejemplo 46. Las caras laterales de la nueva pirámide son triángulos isósceles. A menos que sean equiláteros, tomamos uno de ellos como base y repetimos el proceso, disminuyendo de nuevo $S$. Por el principio de variación parcial (ejemplo 22), si hay un tetraedro con un $V$ dado y una $S$ mínima, éste debe ser el tetraedro regular.

**51.** Ver el ejemplo 53.

**52.** Ver el ejemplo 53.

**53.** $V$, $S$ e $y$ denotan el volumen, la superficie y la altura de la pirámide, y $x$, $a$ y $l$ están conectados con la base de la pirámide del mismo modo que en la solución del ejemplo 36 con la base del prisma. Con $p$ denominamos el radio del círculo inscrito en la base. Luego,

$$\begin{aligned} V &= ax^2y/3 \\ ax^2 &= lxp/2 \\ S &= ax^2 + lx(p^2 + y^2)^{1/2}/2 \\ &= ax^2 + (4a^2x^4 + l^2x^2y^2)^{1/2}/2. \end{aligned}$$

Intentado introducir expresiones que dependen de la forma, aunque no del tamaño, nos vemos conducidos a considerar

$$\frac{S}{ax^2} = 1 + \left[1 + \left(\frac{ly}{2ax}\right)^2\right]^{1/2} = 1 + (1 + t)^{1/2}$$

(introducimos una abreviación si hacemos $[ly/(2ax)]^2 = t$) y

$$\frac{S^2}{(3V)^2} = \frac{l^2}{4a} \frac{[1 + (1 + t)^{1/2}]^3}{t}.$$

Como $V$ está dado y $S$ debe ser el mínimo, el lado izquierdo debe ser un mínimo. En consecuencia, el lado derecho será un mínimo. Sin embargo, la forma está dada y por ello $l^2/a$ está dado. Por tanto, todo lo que queda es encontrar el valor de $t$ que hace del lado derecho un mínimo: este valor es *independiente de la forma.* Sirve igualmente para todas las formas especiales, por ejemplo, las mencionadas en el ejemplo 51 y 52. Sin embargo, hay una forma especial, para la cual sabemos el resultado: si la base es un triángulo equilátero, la mejor razón $S : ax^2$, o la superficie total a la base, es 4 : 1, por el ejemplo 50. Esto resulta cierto para todas las formas, puesto que $S/(ax^2)$ depende sólo de $t$, y da

$$1 + (1 + t)^{1/2} = 4, \quad t = 8.$$

**54.** El lector copiará la siguiente tabla, sustituyendo para el número de cada problema una figura conveniente.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| (*a*) | 34 | 35 | 36 | 42 | *u* | 48 | *x* |
| (*b*) | 51 | 52 | 53 | 46 | 50 | *v* | *y* |
| (*c*) | 37 | 38 | 39 | 47 | *w* | 49 | *z* |
| (*d*) | — | — | — | — | 40 | 41 | *n* |

Las filas: (*a*) prismas, (*b*) pirámides, (*c*) dobles pirámides, (*d*) polígonos (relevante sólo para las últimas tres columnas).

Las columnas: (1) recto con base cuadrada, (2) recto con círculo como base, (3) recto con base de forma dada, (4) transición del oblicuo al recto, (5) base triangular arbitraria, (6) base cuadrilateral arbitraria, (7) base poligonal arbitraria con un número dado $n$ de lados.

La analogía puede sugerir los teoremas que se espera que llenen los abismos señalados por las letras $u$, $v$, $w$, $x$, $y$, $z$ y $n$. He aquí algunos:

(*u*) El prisma que, de todos los prismas triangulares con un volumen dado, tiene la superficie mínima posee las siguientes propiedades: su base es un triángulo equilátero, el área de su base es 1/6 del del área total de su superficie, está circunscrito a una esfera que toca cada cara en su centro.

(*y*) La pirámide que, de todas las pirámides con base poligonal de $n$ lados y un volumen dado tiene la superficie mínima, posee las siguientes propiedades: su base es un polígono regular, el área de su base es 1/4 del área total de su superficie, está circunscrita a una esfera que toca cada cara en el centroide de ésta.

Sobre la base de lo anterior podemos probar fácilmente (*u*), (*v*) y

$(w)$, pero $(x)$, $(y)$ y $(z)$ dependen de $(n)$, que discutiremos más tarde; ver la sección 10.7 (1).

**55.** (1) Empleando el *método* y la notación de la sección 6 tenemos

$$V = abc, \qquad S_5 = ab + 2ac + 2bc.$$
$$(S_5/3)^3 \geqq ab \cdot 2ac \cdot 2bc = 4V^2$$

con igualdad si, y sólo si,

$$ab = 2ac = 2bc$$

o $a = b = 2c$: la caja es la mitad de un cubo.

(2) Empleando el *resultado* de la sección 6 consideramos el plano de la cara no contada en $S_5$ como un espejo. La caja junto con su imagen del espejo forma una nueva caja del volumen $2V$ cuya cuperficie *total* está dada, $= 2S_5$. Por la sección 6 la nueva caja (doble) será un cubo cuando se alcanza el máximo del volumen.

**56.** Siguiendo el ejemplo 55 consideremos el plano de la cara oculta como un espejo. El máximo para la pirámide triangular que es la mitad de un cubo está dividida por un plano diagonal. Aplicar el ejemplo 48, con una observación adicional.

**57.** [Putnam, 1950]. Siguiendo los ejemplos 55 y 56 consideremos los planos de las caras ocultas como espejos. El máximo para la pirámide triangular es un cuerpo del cubo; el cubo está dividido en cuatro fragmentos congruentes por dos planos de simetría; un plano diagonal, un plano perpendicular al primero y paralelo a dos caras.

**58.** $A$, $L$, $r$ y $s$ denotan el área, el perímetro, el radio y el arco del sector, respectivamente. Luego,

$$A = rs/2, \qquad L = 2r + s$$
$$(L/2)^2 \geqq 2r \cdot s = 4A;$$

empleamos el teorema de las medias. La igualdad se alcanza cuando $s = 2r$ y el ángulo del sector igual a dos radiantes.

**59.** Sean $u$, $v$ y $w$ los lados del triángulo, $A$ el área, $y$ el ángulo dado opuesto a $w$. Luego,

$$2A = uv \operatorname{sen} y.$$
$$(1) \qquad [(u + v)/2]^2 \geqq uv = 2A/\operatorname{sen}\ y,$$

por el teorema de las medias. La igualdad se alcanza cuando $u = v$, es decir, el triángulo es isósceles.

(2)
$$\begin{aligned} w^2 &= u^2 + v^2 - 2uv\cos y \\ &= u^2 + v^2 - 4A\cot y \\ (u^2 + v^2)/2 &\geqq uv = 2A/\operatorname{sen} y. \end{aligned}$$

La igualdad se alcanza, y por ello $w$ es un mínimo, cuando $u^2 = v^2$ y el triángulo es isósceles.

(3) Tanto $u + v$ y $w$, como también $u + v + w$, es un mínimo, cuando el triángulo es isósceles.

**60.** Usamos la notación del ejemplo 59. El punto dado está sobre el lado $w$. Trazar desde el punto dado paralelas a $u$ y $v$, terminando en $v$ y $u$, a las que llamaremos $a$ y $b$, respectivamente; a $b$ y $a$ son dados (son, en efecto, coordinadas oblicuas). A partir de triángulos semejantes

$$\frac{v - b}{a} = \frac{b}{u - a} \text{ or } \frac{a}{u} + \frac{b}{v} = 1$$

$$\frac{1}{4} = \left(\left[\frac{a}{u} + \frac{b}{v}\right]/2\right)^2 \geqq \frac{ab}{uv} = \frac{ab\ \operatorname{sen}\ y}{2A}$$

$$A \geqq 2ab\operatorname{sen} y.$$

Hay igualdad si, y sólo si,

$$\frac{a}{u} = \frac{b}{v} = \frac{1}{2}, \qquad u = 2a, \qquad v = 2b$$

y el punto dado es el punto medio de $w$.

**61.** Usemos la notación del ejemplo 6 y el teorema de las medias.

(1)
$$\begin{aligned} V &= abc \leqq [(a + b + c)/3]^3 = [E/12]^3 \\ S &= 2ab + 2ac + 2bc \\ &\leqq a^2 + b^2 + a^2 + c^2 + b^2 + c^2 \\ &= 2(a + b + c)^2 - 4(ab + ac + bc), \end{aligned}$$

es decir:

$$3S \leqq 2(E/4)^2.$$

En ambos casos, la igualdad se alcanza sólo para $a = b = c$, es decir, para el cubo.

**62.** Usemos la notación de la sección 6 y el teorema de las medias. La longitud es $c$, el contorno $2(a + b)$ y

$$V = (2a \cdot 2b \cdot c)/4 \leqq [(2a + 2b + c)/3]^3 /4 \leqq l^3/108.$$

La igualdad se alcanza sólo para $2a = 2b = c = l/3$.

**63.** Usemos la notación de la solución del ejemplo 35. Luego,

$$d^2 = (2x)^2 + (y/2)^2 = 2(x^2 + x^2 + y^2/8)$$

y, en consecuencia, por el teorema de las medias,

$$V^2 = \pi^2 x^4 y^2 = 8\pi^2 x^2 \cdot x^2 \cdot y^2/8 \leqq 8\pi^2 (d^2/6)^3,$$

con igualdad sólo si

$$x^2 = y^2/8 = d^2/6.$$

Para el trasfondo histórico cf. O. TOEPLITZ, *Die Entwicklung der Infinitesimalrechmung*, págs. 78-79.

Sin solución: **23**, **33**.

## CAPITULO IX

**1.** (1) Imaginemos dos espejos perpendiculares al plano del dibujo, uno que pasa por $l$ y otro por $m$. Una persona en $P$ mira $m$ y se ve a sí mismo al otro lado: la luz que viene de $P$ retorna a $P$ tras una primera reflexión en $l$ y una segunda en $m$. La luz, que elige la trayectoria más corta, describe el deseado $\Delta PYZ$ con perímetro mínimo; los lados de $\Delta PYZ$ comprenden ángulos iguales con $l$ y $m$ en los puntos $Y$ y $Z$, respectivamente. (2) Sean $P'$ y $P''$ imágenes en el espejo de $P$ con respecto a $l$ y $m$. La línea recta que une $P'$ y $P''$ intersecta $l$ y $m$ en los requeridos $Y$ y $Z$, respectivamente, y su longitud es la del perímetro mínimo deseado. (Por la idea de la figura 9.3, aplicada dos veces.)

**2.** (1) La luz, tras tres reflexiones sucesivas en tres espejos circulares, retorna a su origen desde la dirección opuesta.

(2) Una cinta de goma une tres anillos rígidos. Las dos interpretaciones sugieren que los dos lados del triángulo requerido que se reúnen en un vértice sobre un círculo dado forman ángulos iguales con el radio.

**3.** Un «boomerang» de luz o una cinta de goma, como en el ejemplo 2; $XY$ y $XZ$ están igualmente inclinados a $BC$, etc.

**4.** Un polígono de $n$ lados y perímetro mínimo, inscrito en un polígono dado de $n$ lados, tiene la siguiente propiedad: aquellos dos lados del polígono mínimo, cuyo vértice común está en cierto lado $s$ del polígono dado, están inclinados igualmente hacia $s$. Ver, no obstante, los ejemplos 6 y 13.

**5.** Llamemos $A$ al punto de intersección de $l$ y $m$. Tomemos un punto $B$ en $m$ y otro punto $C$ en $l$, de manera que $\angle BAC$ (menor de 180°) contiene el punto $P$ en su interior. Luego, por reflexión,

$$\angle P''AB = \angle BAP, \qquad \angle PAC = \angle CAP'$$

y de aquí

$$\angle P''AP' = 2 \angle BAC.$$

La solución falla cuando $\angle P''AP' \geqq 180°$ o, lo que es lo mismo, cuando el ángulo dado $\angle BAC \geqq 90°$.

**6.** La solución no tiene aplicación cuando el triángulo dado tiene un ángulo $\geqq 90°$; ver el ejemplo 5. La solución del ejemplo 4 es, naturalmente, confiable *a fortiori*, salvo excepción.

**7.** Fijar de momento $X$ en la posición $P$ sobre el lado $BC$. Luego la solución (2) del ejemplo 1 se puede aplicar (puesto que $\angle BAC$ es agudo, ver el ejemplo 5); el perímetro mínimo es $P'P''$. Ahora bien, la longitud $P'P''$ depende de $P$; queda por encontrar el mínimo de $P'P''$. (Como $P'P''$ mismo fue obtenido como mínimo, buscamos el mínimo de los mínimos o «minimun minimorum»). Ahora, por reflexión, $P''A = PA = P'A$. Por tanto, $\Delta P''AP'$ es isósceles; su ángulo en $A$ es independiente de $P$ (ver el ejemplo 5) y, por ello, su *forma es independiente de P*. En consecuencia, $P'P''$ es mínimo cuando $P'A = PA$ es mínimo, y éste es visiblemente el caso cuando $PA \perp BC$; cf. las sección 8.3. *Los vértices del triángulo con perímetro mínimo inscrito en un triángulo agudo dado son los pies de las tres alturas del triángulo dado.* Comparando esto con la solución del ejemplo 3 vemos que *las alturas de un triángulo agudo bisecciona los ángulos respectivos del triángulo cuyos pies son los vértices.* El último resultado es, por supuesto, bastante elemental. La presente solución se debe a L. Fejér. Cf. COURANT-ROBBINS, págs. 346-353.

**8.** No. Si $\Delta ABC$ tiene un ángulo $\geqq 120°$, el vértice de ese ángulo es el centro de tráfico. Esto es fuertemente sugerido por la solución mecánica de la sección 2 (2).

**9.** [Putnam, 1949]. Esto es muy análogo al problema del plano más sencillo tratado en la sección 1 (4), la sección 2 (2) y el ejemplo 8.

¿Qué métodos debemos adoptar? (1) La *interpretación mecánica*, modelada en la figura 9.7. Hay cuatro poleas, una en cada uno de los cuatro puntos dados $A$, $B$, $C$ y $D$. Cuatro cuerdas están atadas juntas en el punto $X$; cada cuerda pasa por una de las poleas y soporta un peso de una libra en el extremo. Como en la sección 2 (2), una primera consideración (de la energía potencial) muestra que la posición de equilibrio de este sistema mecánico corresponde al propuesto problema de mínimos. Una segunda consideración se refiere a las fuerzas que actúan sobre el punto $X$. Hay cuatro fuerzas tales; son iguales en magnitud y en la dirección de las cuatro cuerdas tensas van a $A$, $B$, $C$ y $D$, respectivamente. La resultante de las dos primeras fuerzas deben contrabalancear la resultante de las dos últimas fuerzas. En consecuencia, estas resultantes están en la misma línea que biseca $\angle AXB$ y $\angle CXD$. La igualdad de estos ángulos se sigue de la congruencia de dos paralelogramos de fuerzas (ambos son rombos). De modo semejante están relacionados los siguientes pares: $\angle AXC$ y $\angle BXD$, $\angle AXD$ y $\angle BXC$. (2) *Variación parcial e interpretación óptica*, modelada en la figura 9.4. Mantengamos constante (por un momento) $CX + DX$, la suma de dos distancias. Luego el punto $X$ tiene que variar en la superficie de un esferoide oblongo (elipsoide de revolución) con focos en $C$ y $D$. Concebimos esta superficie como un espejo. La luz, que saliendo de $A$ se refleja en nuestro espejo esferoide y llega a $B$, hace a $AX + XB$ un mínimo; a lo largo de su trayectoria $\angle AXB$ es cortado por la normal al espejo en el punto $X$. La misma normal corta el $\angle CXD$ por la sección 1 (3) o por la sección 2 (1). No obstante, la igualdad del $\angle AXB$ y del $\angle CXD$ no se obtiene tan fácilmente por este método: aunque ambos métodos resultan bien en el caso análogo más sencillo, su aplicación al presente teorema no es tan oportuna.

**10.** Sí. Dondequiera que esté el punto $X$,

$$AX + XC \geqq AC, \qquad BX + XD \geqq BD,$$

puesto que la línea recta es la más corta entre dos puntos, y ambas desigualdades se convierten en ecuaciones solamente si $X$ es el punto de intersección de las diagonales $AC$ y $BD$: esto es, el centro de tráfico. El enunciado del ejemplo 9 continúa siendo totalmente correcto a la vista del hecho de que la normal al plano del cuadrilátero es un bisector común al $\angle AXC$ y al $\angle BXD$, que son ambos ángulos rectos.

**11.** Se sigue con variación parcial del resultado de la sección 1 (4). Cf. COURANT-ROBBINS, págs. 354-361.

**13.** Llevar la bola paralelamente a una diagonal de la mesa. La figura 9.14 se aplica al ejemplo 12 cuatro veces sucesivamente. Imaginar la figura 9.14 dibujada en papel transparente y doblarla a lo largo de las líneas de reflexión; entonces, los diversos segmentos de la línea recta $PP$ cubren exactamente la trayectoria rómbica de la bola de billar. Y, a propósito, obsérvese aquí un caso en que el ejemplo 4 tiene una infinidad de soluciones.

**14.** Por la figura 9.15 (la cual será dibujada en un papel transparente),

$$n2\alpha < 180° < (n + 1)\,2\alpha,$$

$$90°/(n + 1) < \alpha < 90°/n.$$

Dibujar figuras que ilustren los casos $n = 1, 2, 3$. Considerar el caso $n = \infty$.

**15.** El caso particular tratado en la sección 1 y el ejemplo 12 proporcionan varias sugerencias; ver ejemplos 16, 17 y 18.

**16.** Si $A$, $B$ y $AX + XB$ nos son dados, el *locus* de $X$ es la superficie de un esferoide oblongo (elipsoide de revolución) con focos $A$ y $B$; tales esferoides son las superficies de nivel. El esferoide al cual la línea dada $l$ es tangente al punto $X$ de la solución. La normal a este esferoide en el punto $X$ es *perpendicular* a $l$, y bisecciona al $\angle AXB$ por una propiedad de la elipse probada en la sección 1 (3) y en la sección 2 (1).

**17.** Colocar una hoja de papel doblado en dos de modo que el doblez coincida con $l$ y una mitad de la hoja pase por $A$ y la otra mitad pase por $B$. La línea más corta deseada está descrita sobre esta hoja doblada. Si la hoja no está doblada, la línea más corta se hace recta. Tanto en la hoja doblada como en la no doblada las líneas $XA$ y $XB$ *comprenden con l el ángulo*.

**18.** Fijar un cabo de una cinta de goma de longitud adecuada a $A$, pasar la cinta por la vara rígida $l$ en $X$ y fijar el otro cabo a $B$ de modo que la cinta está tirante: así se forma la línea más corta requerida por el ejemplo 15 (si la fricción es desdeñable). Tres fuerzas actúan en el punto $X$: dos tensiones iguales en magnitud, una dirigida hacia $A$ y otra hacia $B$, y la reacción de la vara que es perpendicular a $l$ (supuesto que la fricción es desdeñable). El paralelogramo de fuerzas es un rombo, y así, una normal a $l$ bisecciona $< AXB$, como vimos en el ejemplo 16. La reacción de la vara carece de componente paralelo a la vara, y así, los componentes de las tensiones paralelas a $l$ deben ser iguales en cantidad (y opuestas en dirección). Por tanto, $XA$ y $XB$ están igualmente

inclinados a *l*, como vimos en el ejemplo 17. A propósito, la equivalencia de resultados del ejemplo 16 y el 17 puede ser mostrado por un poco de geometría sólida. (Los ángulos triedros son congruentes si tienen tres datos apropiados en común.)

**19.** Una cinta de goma se mantiene rígidamente alrededor de tres agujas de media. Variación parcial y el ejemplo 16, o el 18; las bisectrices de los tres ángulos de un triángulo se juntan en el centro del círculo inscrito. El ejemplo 3 es un caso límite.

**20.** Cada vértice del triángulo es el punto medio de una arista del cubo. El triángulo es equilátero; su centro es el centro del cubo; su perímetro es $3\sqrt{6}a$.

**21.** Por variación parcial, la sección 8.3 y la sección 1 (4) o 2 (2), $TX$, $TY$ y $TZ$ son perpendiculares a $a$, $b$ y $c$, respectivamente, e igualmente inclinados uno a otro (120°). Podemos llamar $T$ el «centro de tráfico de tres líneas oblicuas». El problema de la sección 1 (4) es un caso extremo: $a$, $b$ y $c$ se hacen paralelos. Hay una generalización obvia y hay algunos problemas obvios análogos: el centro de tráfico de tres esferas, el centro de tráfico de un punto, una línea recta y un plano, y así sucesivamente.

**22.** El centro de tráfico de tres aristas oblicuas de un cubo es, por supuesto, el centro del cubo. Representa claramente la rotación del cubo en 120° que intercambia las tres aristas oblicuas dadas, y la situación del triángulo encontrado en el ejemplo 20.

**23.** Con objeto de encontrar la línea más corta entre dos puntos dados $A$ y $B$ sobre la superficie de un poliedro, imaginemos la superficie poliédrica hecha de cartón, de polígonos planos enlazados entre sí y adecuadamente doblados. Desdoblamos la superficie poliédrica en un plano: la línea más corta buscada resulta ser la *línea recta de A a B*. No obstante, antes de desdoblarlo tenemos que cortar la superficie poliédrica a lo largo de aristas convenientes cuyas líneas más cortas *no se cruzan*. Como no sabemos por adelantado qué caras y aristas cruzarán la línea más corta, tenemos que examinar todas las combinaciones posibles. Volvamos ahora al problema propuesto, leamos las secuencias esenciales de las caras y observemos tras cada secuencia el cuadrado de la distancia rectilínea desde la araña a la mosca a lo largo de la secuencia.

(1) Pared final, techo, pared final:

$$(1 + 20 + 7)^2 = 784;$$

(2) Pared final, techo, pared lateral, pared final:

$$(1 + 20 + 4)^2 + (4 + 7)^2 = 746;$$

(3) Pared final, techo, pared lateral, suelo, pared final:

$$(1 + 20 + 1)^2 + (4 + 8 + 4)^2 = 740.$$

**24.** Un arco de un gran círculo es una geodésica sobre la esfera. Un gran círculo es una curva del plano; el plano en que el gran círculo yace es un plano osculatriz en todos los puntos. Este plano pasa por el centro de la esfera, y, por tano, contiene todas las normales a la esfera (todos los radios) que pasan por los puntos del gran círculo. Un pequeño círculo no es una geodésica; en efecto, el plano del pequeño círculo no contiene ninguna de las normales a la esfera que pasa por los puntos del pequeño círculo.

**25.** Por la conservación de la energía, la magnitud de la velocidad del punto es constante, aunque, por supuesto, la dirección de la velocidad varía. La diferencia de los sectores de velocidad en los dos puntos finales de un corto arco de la trayectoria se debe a las normales reacciones de la superficie y es, por tanto, casi normal a la superficie. Esta es la propiedad característica de una geodésica; ver el ejemplo 24 (2). Otra versión del mismo argumento: reinterpretar las tensiones a través de la cinta de goma del ejemplo 24 (2) como velocidades a lo largo de la trayectoria; todos los vectores son de la misma magnitud y la variación en la dirección se debe a reacciones normales en ambos casos.

**26.** Empujar las $n$ aristas libres contra un plano (su pupitre), formando una pirámide con $n$ caras laterales isósceles cuya base es el polígono deseado. En efecto, la base está inscrita en un círculo, cuyo centro es el pie de la altura de la pirámide. El radio del círculo es el tercer lado de un triángulo rectángulo cuya hipotenusa es el radio del gran círculo trazado sobre cartulinas y el segundo lado la altura de la pirámide.

**27.** Si el centro de gravedad está tan próximo al suelo como sea posible, hay equilibrio. Tan poca mecánica como ésta es bastante para sugerir la solución buscada: tomemos un punto $D$ sobre la superficie de $P$ tal que la distancia $CD$ sea un mínimo. Un sencillo examen muestra que $D$ no puede ser un vértice de $P$ ni estar sobre la arista de $P$, y que $CD$ es perpendicular a la cara $F$ de P sobre la que $D$ está. Ver Polya-Szegö, *Analysis*, vol. 2, pág. 162, problema 1.

**28.** (*a*) Imaginar el globo completamente seco, de modo que los

picos, valles y profundidades quedan expuestos. Ahora, cubrir sólo una de las profundidades con agua. La parte restante del globo tiene $P$ picos, $S$ pasos y $D$—1 profundidades, pudiendo ser considerada como una isla. Por el resultado probado en el texto,

$$P + (D - 1) = S + 1.$$

(*b*) Las líneas de nivel y las líneas de descenso subdividen el globo en $C$ «países»; ésta es la terminología del ejemplo 3.2. Tomemos tantas líneas de cada punto observable —pico, profundidad o paso que alcanzan un vértice— como en las figuras 9.16 y 9.17, y de modo que ningún país tenga más de un punto observable en sus límites.

Distribuimos cada arista, o línea límite, de modo igual entre los países que separa, dando 1/2 de la arista a cada país. De modo semejante distribuimos cada vértice igualmente entre los países del cual es un vértice. En cambio, cada país contribuirá al lado izquierdo de la ecuación de Euler

$$V - A + C = 2;$$

contribuirá una unidad a $C$ y una fracción adecuada a $V$ y a $-A$. Calculemos esta contribución para las varias clases de países.

I. Si no hay punto digno de nota en sus límites, el país es un cuadrilátero, contenido entre dos líneas de nivel y dos líneas de descenso. Su contribución a $V - A + C$ es

$$4 \times \frac{1}{4} - 4 \times \frac{1}{2} + 1 = 0.$$

II. Si hay un pico o una profundidad en sus límites, el país es un triángulo; ver la figura 9.16. Si el pico o la profundidad es un vértice común a $n$ países, la contribución de cada país a $V - A + C$ es

$$\left(2 \times \frac{1}{4} + \frac{1}{n}\right) - 3 \times \frac{1}{2} + 1 = \frac{1}{n}$$

y la contribución conjunta de todos los $n$ países es $n \cdot 1/n = 1$.

III. Si hay un paso en sus límites, el país es un cuadrilátero; ver la figura 9.17. Su contribución a $V - A + C$ es

$$\left(3 \times \frac{1}{4} + \frac{1}{8}\right) - 4 \times \frac{1}{2} + 1 = -\frac{1}{8}$$

y la contribución conjunta de los ocho países cuyo paso es el vértice común es $8 \cdot (-1/8) = -1$.

El total definitivo de todas las contribuciones es, por el teorema de Euler,

$$P + D - S = 2.$$

(*c*) La prueba que usa la idea del «Diluvio» no es propiamente un ejemplo de «matemática física»: usa las ideas en contacto con la experiencia diaria, pero no bajo ninguna específica teoría física. La indicación de la parte (*b*) de la cuestión fue algo desorientadora: parecía sugerir que $P$, $D$ y $S$ son de algún modo análogo a $C$, $V$ y $A$, lo cual no es así de ningún modo. Aun así, fue una indicación útil: llamó nuestra atención sobre el teorema de Euler. Esto es, no obstante, bastante natural: las ideas que nos guían en la solución de problemas están a menudo equivocadas, aunque aun así sean útiles.

**29.** Sea $t$ el tiempo de descenso de la piedra y $t_2$ el del ascenso del sonido. Luego,

$$t = t_1 + t_2, \quad d = gt_1^2/2, \qquad d = ct_2.$$

Eliminando $t_1$ y $t_2$, y resolviendo una ecuación cuadrática, encontramos

$$d^{1/2} = -c(2g)^{-1/2} \pm [c^2(2g)^{-1} + ct]^{1/2}.$$

Puesto que $t = 0$ daría $d = 0$, tenemos que elegir el signo $+$. Luego,

$$d = c^2g^{-1} + ct - c^2g^{-1}[1 + 2gc^{-1}t]^{1/2}.$$

(b) $$d = gt^2/2 - g^2t^3/(2c) + \ldots$$

Desdeñando los términos no escritos aquí, podemos usar los dos términos retenidos como una fórmula aproximada muy adecuada.

(*c*) Es típico que podamos prever el término principal de la extensión y también el signo de la corrección sobre la base de consideraciones físicas. También el procedimiento matemático usado para obtener una adecuada fórmula aproximada es típico: *desarrollmos* (la expresión para *d*) *en valores de una pequeña cantidad* (el tiempo *t*). Cf. la sección 5.2.

**30.** El espejo elíptico se convierte en espejo parabólico que recoge todos los rayos de luz que caen paralelos a su eje dentro del foco. Un espejo parabólico tal es la parte más esencial de un telescopio reflexivo.

**31.** La ecuación es separable. Obtenemos

$$dx = \left(\frac{y}{c - y}\right)^{1/2} dy$$

por transformaciones obvias. Pongamos

$$\left(\frac{y}{c - y}\right)^{1/2} = \tan \varphi,$$

introduciendo la variable auxiliar $\varphi$, y obtenemos

$$y = c \operatorname{sen}^2 \varphi, \qquad x = c(\varphi - (1/2) \operatorname{sen} 2\varphi).$$

Encontramos $x$ por integración y tenemos que elegir la constante de integración de modo que la curva pase por el origen: $\varphi = 0$ implica $x = y = 0$. Poniendo $2\varphi = t$, $c = 2a$, obtenemos las ecuaciones usuales de la cicloide:

$$x = a(t - \operatorname{sen} t), \qquad y = a(1 - \cos t).$$

La cicloide pasa por el punto $A$, que es el origen. Hay sólo un valor de $a$ que hace la primera rama de la cicloide (correspondiente a $0 < t < 2\pi$) y pasa por el punto $B$. Al objeto de ver esto, $a$ varía de 0 a $\infty$; esto «infla» la cicloide que, barriendo un cuadrante del plano, toca $B$ cuando alcanza el tamaño recto.

**33.** Sea $a$ el radio de la esfera (como en la sección 5), $h$ la altura del segmento, $V$ su volumen y $C$ el volumen del cono con la misma base del segmento y la misma altura $h$. El origen (el punto 0 en la figura 9.13) es el vértice común del segmento y el cono. Desde la geometría elemental y la ecuación del círculo dado en la sección 5,

$$C = \frac{\pi(2ah - h^2)h}{3}$$

Utilizar la figura 9.13, pero considerar ahora sólo las secciones transversales a la distancia $x$ desde 0 con $0 < x < h$. Pasando del equilibrio de las secciones transversales expresadas por la ecuación (A) al equilibrio de los sólidos (segmento, cono —aunque no con volumen $C$— y cilindro) encontramos

$$\text{(B)} \qquad 2a(V + \pi h^2 \cdot h/3) = (h/2)\pi(2a)^2 h.$$

De aquí,

$$V = \frac{\pi h^2(3a - h)}{3} = \frac{a + (2a - h)}{2a - h} C;$$

$2a - h$ es la altura del segmento complementario.

**34.** Escribir la ecuación del círculo considerada en la sección 5 en la forma

(A) $$2a\pi x^2 = x\pi y^2 + x\pi x^2.$$

Sólo $\pi x^2$, la sección transversal de un cono, cuelga ahora del punto $H$ de la figura 9.13; la sección transversal $\pi y^2$ de la esfera y la sección transversal $\pi x^2$ de otro cono (congruente con el primero) queda en su posición original (con la abscisa $x$). Considerar $0 < x < a$, pasar el equilibrio de los tres sólidos, introducir $x$, la abscisa del centro de gravedad de la hemisfera y recordar la posición del centro de gravedad de un cono (su distancia desde el vértice es 3/4 de la altura):

(B) $$2a \cdot \pi a^2 \cdot a/3 = \bar{x} \cdot 2\pi a^2/3 + (3a/4)\pi a^2 \cdot a/3,$$

$$\bar{x} = 5a/8.$$

**35.** Mantener la notación del ejemplo 33, pero cambiar la del ejemplo 34 en un respecto: $x$ denota ahora la abscisa del centro de gravedad de un segmento con altura $h$. Considerando $0 + x + h$, pasar desde (A) del ejemplo 34 a

(B) $$2a \cdot \pi h^2 \cdot h/3 = \bar{x}V + (3h/4)\,\pi h^2 \cdot h/3,$$

que da, a la vista del valor de $V$ encontrado en el ejemplo 33,

$$\frac{\bar{x}}{h - \bar{x}} = \frac{h + 4(2a - h)}{h + 2(2a - h)}.$$

**36.** $h$ denota la altura y $V$ el volumen del segmento. Escribir la ecuación usual de la parábola en la forma

(A) $$2p \cdot \pi y^2 = x \cdot \pi(2p)^2.$$

Observar la sección transversal $\pi y^2$ de la paraboloide y la sección transversal $\pi(2p)^2$ de un cilindro. Considerando $0 < x < g$, y, pasando de las secciones transversales a la de los sólidos, encontramos:

(B) $$2p \cdot V = (h/2)\pi(2p)^2 h,$$

$$V = \pi p h^2 = (3/2)\pi 2ph(h/3).$$

Observar que, por la ecuación de la parábola, $\pi 2ph$ es la base del segmento.

**37.** Mantener la notación del ejemplo 36 y $x$ que denota la abscisa del centro de gravedad del segmento. Ahora, escribir la ecuación de la parábola en la forma

$$x \cdot \pi y^2 = 2p \cdot \pi x^2. \tag{A}$$

Notar $\pi x^2$, la sección transversal de un cono. Considerando $0 < x < h$ y pasando de las secciones transversales a los sólidos, encontramos

$$\bar{x}V = 2p \cdot \pi h^2(h/3),$$

y de aquí, por el ejemplo 36,

$$\bar{x} = 2h/3.$$

**38.** $n = 0$: volumen del prisma, área del paralelogramo; $n = 1$: área del triángulo, centro de gravedad del paralelogramo o prisma: $n = 2$: volumen del cono o pirámide, centro de gravedad del triángulo; $n = 3$: centro de gravedad del cono o pirámide.

Observar que el método de Arquímedes, como ha sido presentado aquí en la sección 5 y los ejemplos 33-38, sería muy adecuado para una clase de geometría analítica y podría proporcionar nuevo interés por este tema, que suele resultar árido y aburrido en las presentaciones usuales. Las proposiciones del *Método* que no hemos discutido pueden ser tratadas de modo semejante y usadas de una manera muy parecida.

Sin solución: **12, 32.**

## CAPITULO X

**1.** No. El abismo no es demasiado malo: la existencia del máximo puede ser establecida con la ayuda del teorema general citado en el capítulo VIII, nota 3, a pie de página.

**2.** La fórmula explícita dada en la solución del ejemplo 8.41 muestra que $A^2 \leqq (p-a)(p-b)(p-c)(p-d)$, y la igualdad se obtiene si, y sólo si, $\varepsilon = 180°$, en cuyo caso el cuadrilátero se inscribe en un círculo.

**3.** Sean $A$, $B$ y $C$ vértices consecutivos del polígono regular con $n$ lados, y $M$ el punto medio del lado $BC$. Remplazar $\Delta ABM$ por el triángulo isósceles $\Delta AB'M$ (en el que $AB' = B'M$), que tiene la misma

base $AM$ y el mismo perímetro, y, por tanto, un área mayor; ver el ejemplo 8.8.

**4.** Si expresamos las dos áreas en términos de $r$, el radio del círculo, y $n$, el número de lados del polígono, queda por probar la desigualdad:

$$\frac{\pi^2 r^2}{n \tan(\pi/n)} < \pi r^2.$$

Es más elegante observar que un polígono regular es circunscribible a un círculo: el resultado deseado es un caso particular del ejemplo 5.

**5.** Un polígono con área $A$ y perímetro $L$ se circunscribe a un círculo con radio $r$. Luego, obviamente, $\pi r^2 < A$. Las líneas trazadas desde el centro del círculo a los vértices del polígono lo divide en triángulos con la altura común $r$; de aquí, $A = Lr/2$. Combinando los dos resultados obtenidos encontramos

$$A = \frac{L^2 r^2}{4A} < \frac{L^2}{4\pi}.$$

Ahora, $L^2/(4\pi)$ es precisamente el área del círculo que tiene el perímetro $L$.

**6.** $A$ denota el área y $L$ el perímetro de una curva dada, y $r$ el radio del círculo con el mismo perímetro, de modo que $L = 2\pi r$. $A_n$ denota el área y $L_n$ el perímetro de un polígono $P_n$ que tiende a la curva dada como $n \to \infty$. Luego, $A_n$ tiende a $A$ y $L_n$ a $L$. Considerar el polígono $P_n'$ que es similar a $P_n$ y tiene el perímetro $L$; el área de $P_n'$ es $A_n (L/L_n)^2$. Puesto que $P_n'$ tiene el mismo perímetro que el círculo de radio $r$, concluimos de la sección 7 (4) que

$$A_n (L/L_n)^2 < \pi r^2.$$

Pasando al límite encontramos que

$$\lim_{n \to \infty} A_n (L/L_n)^2 = A \leqq \pi r^2.$$

Esto justifica el enunciado I de la sección 8. Sin embargo, el texto de la sección 7 (5) es objetable: definitivamente *no* probamos que $A < \pi r^2$, como dicho texto parece sugerir. En efecto, la relación expresada por $<$ puede pasar a la expresada por $\leqq$ cuando hemos pasado al límite.

**7.** Ambos enunciados son equivalentes a la desigualdad

$$\frac{216 V^2}{S^2} \leqq 1;$$

$V$ denota el volumen y $S$ el área de la superficie de la caja. En la sección 8.6 probamos esta desigualdad directamente.

**8.** La equivalencia de I′, II″ y III′ es mostrada por el mismo método que la de I, II y III de la sección 8. Sin embargo, I′ no es equivalente a I. En efecto, I′ explícitamente niega la posibilidad dejada en la vaguedad por I de que una curva que no es un círculo pueda tener el mismo perímetro y el mismo área que el círculo. El argumento de la sección 7 (5), tal como ha sido ampliado en el ejemplo 6, prueba I, pero no prueba I′: prueba $\leqq$ que es bastante para I, pero no $<$ que resulta necesario para I′.

**9.** La solución del problema propuesto añadiría a I de la sección 8 lo que es necesario para obtener I′ del ejemplo 8. Referente a los otros puntos ver los ejemplos 10-13.

**10.** Llamemos $C''$ al triángulo más pequeño contenido en $C$, $L''$ su perímetro y $A''$ su área. Luego, obviamente, $L'' < L$ y $A'' > A$. Tomemos $C'$ como un triángulo similar a $C''$ con perímetro $L$; el área de $C'$ es $A' = A''\,(L/L'')^2 > A'' > A$.

**11.** Si $C$ es una curva cualquiera *no convexa* consideremos primero $C''$, la curva menos convexa que contiene a $C$, y luego $C'$, semejante a $C''$, pero con un perímetro igual al de $C$. El argumento total del ejemplo 10, bajo la desigualdad final, puede repetirse en la situación más general.

**12.** Tomemos dos puntos diferentes $P$ y $Q$ sobre la curva cerrada $C'$. Debe haber sobre $C$ un tercer punto $R$ que no pasa por la línea recta entre $P$ y $Q$, puesto que $C$ no puede estar totalmente contenida en una línea recta. Considerar el círculo que pasa por $P$, $Q$ y $R$. Si este círculo no coincide con $C$, hay un punto $S$ sobre $C$ que no está sobre el círculo: el problema del ejemplo 9 es equivalente al del ejemplo 13.

**13.** Si $C$ no es convexo, el ejemplo 11 da la construcción deseada. Si $C$ es convexo, $P$, $Q$ y $S$ son, de alguna manera, los vértices de un cuadrilátero convexo. La región rodeada por $C$ consta de este cuadrilátero y de cuatro segmentos. Cada segmento está limitado por un lado del cuadrilátero y por uno de los cuatro arcos en los que $P$, $Q$, $R$ y $S$ dividen $C$. Siguiendo la idea de Steiner (ver la sección 5 (2), figuras 10.3 y 10.4) consideramos los cuatro segmentos como rígidos (de cartón) y rígidamente ligados a los respectivos lados del cuadrilátero que consideramos como articulados (con junturas flexibles en los cuatro vértices). Adoptamos la notación del ejemplo 8.41. Luego, por nuestra condición principal, $\varepsilon \neq 180°$. Un movimiento ligero del cuadrilátero articulado

cambia $\varepsilon$ en $\varepsilon'$. Elegimos $\varepsilon'$ tan próximo a $\varepsilon$ que los cuatro arcos rígidamente ligados a los lados *todavía forman una curva C' que no intersecta.* Además, elegimos $\varepsilon'$ de modo que

$$|\varepsilon' - 180°| < |\varepsilon - 180°|.$$

Esto implica que el área de $C'$ es más grande que el de $C$, en virtud de la formulada para $A^2$ en la solución del ejemplo 8.41. Sin embargo, $C'$, que consta de los mismos cuatro arcos que $C$, tiene el mismo perímetro.

**14.** Las dos inferencias tienen la misma forma lógica. Sin embargo, la segunda inferencia, que conduce a un resultado obviamente falso, debe ser incorrecta. Por tanto, también la primera inferencia debe ser incorrecta, aunque apunte a un resultado que debe ser cierto. La segunda inferencia es, en efecto, una parodia ingeniosa de la primera, debida a O. Perron.

La diferencia entre los dos casos debe ser alguna circunstancia exterior no mencionada en el texto propuesto. No hay entero más grande. Sin embargo, entre todas las curvas isoperimétricas hay una con el área más grande. Esta, no obstante, no la aprendimos en el ejemplo 10-13.

**15.** La curva $C$ no es un círculo, pero tiene el mismo perímetro que cierto círculo. El área de $C$ no puede ser más grande que el del círculo. De otro modo, habría, como sabemos por el ejemplo 10-13, otra curva $C'$ todavía con el mismo perímetro que el círculo, pero con área mayor, lo cual es imposible en virtud de lo que probamos en el ejemplo 6.

**16.** Dados dos puntos, $A$ y $B$ en la figura 10.13, unidos por una línea recta y una curva variable que comprenden una región. Consideramos la longitud de la curva y el área de la región. En el texto consideramos la longitud comprendida como dada y vimos el máximo del área comprendida. En ambos casos la solución es la misma: un arco de un circulo. Incluso la prueba es esencialmente la misma. Podemos utilizar, tanto aquí como allí, la figura 10.14. Por supuesto, hay diferencias obvias; el segmento (no sombreado) del círculo de la figura 10.14, I está construido ahora desde el área dada, no desde una longitud dada, y utilizamos entonces el teorema II' del ejemplo 8, no el teorema I'.

**17.** Utilizar el ejemplo 16: identificar los puntos $X$ e $Y$ de la figura 10.11 con los puntos $A$ y $B$ de la figura 10.13, respectivamente, y añadir el invariable $\Delta XYZ$ a la figura 10.13. Hay máximo cuando la línea de longitud dada es un arco de círculo.

**18.** En la figura 10.11 consideramos la línea $CY$ como un espejo; sea $X'$ la imagen de $X$ en el espejo, y apliquemos el ejemplo 17 a $\angle XCX'$ y los dos puntos dados $X$ y $X'$ sobre sus dos lados. Hay máximo cuando la línea de longitud dada es un arco de círculo perpendicular a $CY$ en el punto $Y$.

**19.** Utilizar la variación parcial. Considerar $X$ como fijo: la solución es un arco de círculo perpendicular a $CY$, por el ejemplo 18. Considerar $Y$ como fijo: el arco del círculo es también perpendicular a $CX$. Finalmente, la solución es un arco de círculo perpendicular tanto a $CX$ como a $CY$, y de este modo su centro está en $C$, como suponíamos en la sección 9.

**20.** Hay un máximo cuando la línea recta es perpendicular a la bisectriz del ángulo. Esto se seguiría de la simetría, si supiésemos por adelantado que hay sólo una solución. El resultado se sigue, sin tal suposición, del ejemplo 8.59 (2).

**21.** Por la idea de la figura 10.14 hay un máximo cuando los cordeles $BC$ y $DA$ son arcos del *mismo* círculo del cual son cuerdas los palos $AB$ y $CD$.

**22.** Una línea cerrada consiste en $2n$ piezas, $n$ palos alternantes con $n$ cordeles, rodeando un área máxima, cuando todos los palos son cuerdas y todos los cordeles son arcos del mismo círculo.

**23.** Cuando todos los cordeles del ejemplo 22 son de longitud 0, obtenemos la sección 5 (2) y las figuras 10.3 y 10.4.

**24.** Análogos al ejemplo 16: el disco rígido, la superficie variable con área dada y el círculo que forma el borde de ambos corresponde al palo, el cordel y el par de puntos $A$ y $B$, respectivamente. El método del ejemplo 16 es aplicable. (En la figura 10.14 gira el círculo I sobre su diámetro vertical y hacemos lo mismo con el segmento en la base de la figura II, pero cambiamos su parte superior de un modo más arbitrario.) Asumiendo el teorema isoperimétrico en el espacio, obtenemos: el volumen comprendido es un máximo cuando la superficie con área dada es una porción (una zona con una base) de una esfera.

**25.** Tomamos tres planos perpendiculares uno a otro y tomamos como garantía el teorema isoperimétrico en el espacio. Luego, el ángulo triedro alcanza un octante y podemos usar el análogo de la figura 10.12 en el espacio. Por reflexiones sucesivas sobre los tres planos la superficie que corta el octante llega a ser una superficie *cerrada;* su área y el volumen que la rodea es ocho veces el área dada y el volumen cor-

tado por la superficie original, respectivamente. La superficie cerrada con área dada que rodea el máximo volumen es la esfera. Por tanto, en nuestro caso especial del problema propuesto se llega al máximo cuando la superficie con área dada es una porción (1/8) de una esfera con centro en el vértice del ángulo triedro.

**26.** La configuración considerada en la solución del ejemplo 25 es el caso especial $n = 2$ de la siguiente situación general. Hay $n + 1$ planos; $n$ planos que pasan por la misma línea recta y dividen el espacio en $2n$ partes iguales (ángulos diedros) y el último plano es perpendicular a los $n$ anteriores. Estos $n + 1$ planos dividen el espacio en $4n$ ángulos triedros iguales, a cualquiera de los cuales se aplica el método de reflexiones repartidas, usado en la sección 25, que da el mismo resultado: *el volumen cortado es un máximo cuando la superficie de un área dada es una porción de una esfera con centro en el vértice del ángulo triedro.*

(Hay tres configuraciones más que contienen ángulos triedros, al que se aplica el método y da el mismo resultado. Estas configuraciones están conectadas con los sólidos regulares, la primera con el tetraedro, la segunda con el cubo y el octaedro, y la tercera con el dodecaedro y el icosaedro. Su estudio requiere más esfuerzo o más conocimiento preliminar y por ello los enumeramos en la siguiente tabla, que empieza con la sencilla configuración descrita arriba.

| Planos | Partes del espacio | Angulos | | |
|---|---|---|---|---|
| $n + 1$ | $4n$ | 90° | 90° | $180°/n$ |
| 6 | 24 | 90° | 60° | 60° |
| 9 | 48 | 90° | 60° | 45° |
| 15 | 120 | 90° | 60° | 36° |

Los «planos» son planos de simetría, las «partes del espacio» ángulos triedros y los «ángulos» están comprendidos por los tres planos que limitan el ángulo triedro.)

Es natural conjeturar que el resultado permanece válido para cualquier ángulo triedro. Esta conjetura es apoyada inductivamente por los casos enumerados y también por analogía; la conjetura similar obtenida de modo semejante sobre ángulos en un plano (sección 9) ha sido probada (ejemplo 19).

Es también natural extender la conjetura a ángulos poliédricos y ahí podemos encontrar al menos un caso límite accesible a la verificación. Llamamos aquí «cono» a la parte infinita del espacio descrita

por un ángulo agudo que gira sobre uno de sus lados. Buscamos la superficie con área dada que corta el volumen máximo desde el cono. Puede probarse que esta superficie es: (1) una superficie de revolución, (2) una porción de la esfera y, (3) que el centro de la esfera es el vértice del cono. No podemos aquí entrar en detalles, pero debemos observar que la parte (2) de la prueba sale del ejemplo 24 del mismo modo que la solución del ejemplo 17 sale del ejemplo 16. El problema del ejemplo 25, propuesto por Steiner, espera todavía una solución completa.

**27.** Si una región con área $A$ tiene dos bisectores con un punto común, sería dividida por ellos en tres subregiones, dos con área $A/2$ y una tercera con un área que no desaparece, lo cual es obviamente imposible.

**28.** La línea recta es la más corta: $1 < (\pi/2)^{1/2}$.

**29.** Ver el ejemplo 30.

**30.** Asumir que los puntos extremos de un bisector dado están sobre dos lados diferentes que se juntan en el vértice 0, pero ninguno de estos puntos extremos coincide con $O$. Por reflexiones adecuadas (idea de la figura 10.12) obtenemos seis triángulos iguales, uno de los cuales es el triángulo original, y seis arcos iguales, uno de los cuales es el bisector dado. Los seis triángulos forman un hexágono regular con centro $O$. Los seis arcos forman una curva cerrada que rodea una mitad del área del hexágono y, especialmente, el punto $O$, en que se juntan tres ejes de simetría de la curva. Si la longitud del bisector es un mínimo, la curva cerrada debe ser un círculo o un hexágono regular, de acuerdo a cómo todos los bisectores son admitidos (el presente ejemplo 30) o sólo son admitidos bisectores rectos (ejemplo 29); tenemos que usar el teorema II′ del ejemplo 8 o el teorema conjugado con el de la sección 7 (1), respectivamente. La solución del ejemplo 30 es la sexta parte de un círculo con centro en uno de los vértices, la solución del ejemplo 29 es una línea paralela a uno de los lados; en cada caso hay tres soluciones. El bisector dado puede estar en otras situaciones (ambos extremos sobre el mismo lado, o en el mismo vértice, y así sucesivamente, pero la discusión de estas situaciones corrobora el resultado obtenido.

**31.** [Cf. Putnam, 1946]. Sea $O$ el centro del círculo. Si el segmento de la línea recta $PP'$ es bisecado por $O$, llamemos a $P$ y $P'$ *opuestos* uno a otro. Llamamos opuestas a dos curvas si una de ellas consta de puntos opuestos a los puntos de la otra. Sean $A$ y $B$ los extremos del bisector que llamamos abreviadamente $AB$. Sean $A'$, $B'$ y el arco $A'B'$ opuestos a $A$, $B$ y $AB$, respectivamente. Luego $A'B'$ es un bisector.

Sea $P$ un punto común a $AB$ y $A'B'$ (ejemplo 27) y $P'$ opuesto a $P$. Luego también $P'$ es un punto común a $AB$ y $A'B'$. $A$, $P$, $P'$ y $B$ se siguen unos a otros en este orden sobre $AB$ y sean $PB'$ el más corto de los arcos $PA$ y $PB'$. (Esta elección es posible; ésta es, en efecto, una cuestión de notación.) Considerar la curva que consta de dos piezas: el arco $B'P$ (de $A'B'$) y el arco $PB$ (de $AB$). Esta curva es: (1) más corta que $AB$, y (2) más larga que el diámetro $BB'$, que es la trayectoria recta de $B$ a $B'$. Se sigue de (1) y de (2) que $AB$ es más larga que el diámetro $BB'$, y éste es el teorema.

**32.** El eje menor. Ver el ejemplo 33.

**33.** *El bisector más corto de cualquier región es o una línea recta o un arco de un círculo.* Ver el ejemplo 16. *Si la región tiene un centro de simetría* (como el cuadrado, el círculo y la elipse, aunque no el triángulo equilátero), *el bisector más corto es una línea recta.* La prueba es casi la misma que para el círculo (ejemplo 31).

**34.** Prácticamente, el mismo que en el ejemplo 27.

**35.** Ejemplo 16.

**36.** En los cinco casos, el plano del bisector más corto pasa por el centro de la esfera circunscrita.

Tetraedro: cuadrado en un plano paralelo a dos aristas opuestas; tres soluciones.

Cubo: cuadrado paralelo a una de las caras; tres soluciones.

Octaedro: hexágono en un plano paralelo a una de las caras; cuatro soluciones.

Dodecaedro: decágono en un plano paralelo a una de las caras; seis soluciones.

Icosaedro: decágono en un plano perpendicular a un eje que une dos vértices opuestos; seis soluciones.

La prueba es grandemente facilitada en los últimos cuatro casos por una observación general; ver ejemplo 38.

**37.** Sea $O$ el centro de la esfera. Definimos puntos y curvas opuestas como en el ejemplo 31. Sea $b$ un bisector. Luego también $b$ , la curva opuesta a $b$, es un bisector y $b$ y $b'$ tienen un punto común $P$ (ejemplo 34). También $P'$, el punto opuesto a $P$, es un punto común. Los puntos $P$ y $P'$ dividen $b$ en dos arcos, ninguno de los cuales puede ser más corto que la línea más corta que unen $P$ y $P'$, que es la mitad de un círculo grande.

**38.** Cuatro de los cinco sólidos regulares (todos, excepto el tetraedro) tienen un centro de simetría. *Una superficie cerrada con un centro*

*de simetría tiene un bisector que es una geodésica.* La prueba es casi la misma que para la esfera (ejemplo 37).

**39.** (Ver *Elemente der Mathematik*, vol. 4 (1949), página 93, y volumen 5 (1950), pág. 65, problema 65.) Llamar $d$ la distancia del borde del diafragma desde su vértice. El área del diafragma es $\pi d^2$; esta proposición se debe a Arquímedes. Cf. el ejemplo 11.4.

(1) Si el centro de $S$ es el vértice del diafragma, $d = a$, $\pi d^2 = \pi a^2$.

(2) Sea $l$ la línea recta que une el centro $C$ de la esfera $S$ y el centro $C'$ de la otra esfera, de la cual el diafragma es una parte. Sea $A$ la intersección de $l$ con $S$ que está sobre el mismo lado de $C$ como $C'$, y $D$ y $B$ las intersecciones de $l$ con el diafragma y con el plano que pasa por el borde del diafragma, respectivamente. Si el diafragma biseca el volumen de $S$, los puntos $A$, $B$, $C$ y $D$ se siguen el uno al otro *en este orden* a lo largo de $l$. El punto de $l$ más próximo al borde del diafragma es $B$, y $D$ está más lejos de este borde que $C$. Por tanto, $d > a$, $\pi d^2 > \pi a^2$.

(3) Conjetura: Ninguna superficie que biseca el volumen de la esfera con radio $a$ tiene un área menor de $\pi n^2$. La prueba puede ser difícil.

**41.** (1) El máximo $f = n^2$ se alcanza cuando

$$x_1 = x_2 = \ldots = x_n = 1 \quad \text{o} \quad -1.$$

El mínimo $f = 0$ se alcanza por muy diferentes sistemas $x_1, \ldots x_n$ cuando $n \geqq 3$.

(2) El máximo se alcanza como antes únicamente. El mínimo $f = n$, cuando

$$x_1 = n^{1/2}, \quad x_2 = \ldots = xn = 0.$$

y en $n - 1$ casos semejantes.

**42.** La conjetura es correcta para los sólidos regulares con vértices de tres aristas, pero es incorrecta si un vértice tiene más de tres aristas. Ver M. Goldberg, *Tôhoku Mathematical Journal*, vol. 40 (1935), páginas 226-236.

Sin solución: **40**, **43**.

## CAPITULO XI

**1.** (*a*) sí, (*b*) no: $\alpha$ es innecesaria, el área es $ah/2$.

**2.** (*a*) sí, (*b*) no: $\alpha$ y $\beta$ son innecesarias, el área es $mh$.

**3.** $2\pi rh$, *independiente* de $d$. Solución por el método de Arquímedes o por el cálculo integral: de $x^2 + y^2 = r^2$ se sigue

$$y^2\left(\frac{dy}{dx}\right)^2 + y^2 = r^2, \qquad \int_d^{d+h} 2\pi y\left[1 + \left(\frac{dy}{dx}\right)^2\right]^{1/2} dx = 2\pi rh.$$

**4.** Sea $h$ la altura de la zona cuya área se busca. De los triángulos rectángulos semejantes $h : a = a : 2b$, y por ello el área buscada es $2\pi bh = \pi a^2$, *independiente* de $b$. La zona alcanza la esfera completa cuando $b = a/2$ y el círculo cuando $b = \infty$. Cf. ejemplo 10.39.

**5.** ¿Hemos observado la analogía con el ejeplo 1-4? El volumen de la esfera perforada puede obtenerse elementalmente o por el empleo de la geometría analítica y el cálculo integral como

$$\int_{-h/2}^{h/2} \pi y^2 dx - \pi y_1^2 h = \pi h^3/6;$$

$x^2 + y^2 = r^2$ e $y_1$, es la ordenada correspondiente a $x = h/2$.

**6.** $\pi h^3/12$, *independiente* de $a$ y $b$. Solución semejante a la del ejemplo 5, conectada con la del ejemplo 7. En el caso extremo $a = b = 0$ el segmento alcanza una esfera completa con diámetro $h$. Si $h$ es pequeña, la diferencia entre $Mh$ y $V$ se ve intuitivamente que es muy pequeña.

**7.** $\pi c^2 h/6$, *independiente* de $r$. Si $c = h$, el cono degenera en un cilindro y tenemos el caso de la sección 2 y el ejemplo 5. Solución semejante a la del ejemplo 8.

**8.** Con $O$ como origen y $OX$ como eje-$x$, las ecuaciones del círculo y la parábola de la figura 11.3 son

$$(x - d)^2 + y^2 = r^2, \qquad 2px = y^2,$$

respectivamente. Supongamos que $x_1$ y $x_2$ denotan las abscisas de los

puntos de intersección de las dos curvas, $x_1 < x_2$. Entonces, $x_2 - x_1 = h$ y el volúmen buscado es

$$\pi \int_{x_1}^{x_2} [r^2 - (x-d)^2 - 2px]\, dx = \pi \int_{x_1}^{x_2} (x_2 - x)(x - x_1)\, dx = \pi h^2/6,$$

*independiente* de $r$ y $d$; sustituyamos $x - x_1 = t$. Utilizamos la descomposición en factores de un polinomio de segundo grado cuando las dos raíces y el coeficiente de $x^2$ con conocidos.

**9.** (*a*) sí; el volumen es $\pi h^2(a + 2b)/3$. (*b*) no.

**10.** Sí. Como $u_1$ y $u_2$ pueden ser dados arbitrariamente, hay una infinidad de sistemas posibles $u_1, u_2, \ldots u_{10}$ que satisfacen las relación periódica $u_n = u_{n-1} + u_{n-2}$. Examinamos dos sistemas especiales:

$u_1', u_2', u_3', \ldots u_{10}'$ cuando $u_1' = 0, u_2' = 1$

$u_1'', u_2'', u_3'', \ldots u_{10}''$ cuando $u_1'' = 1, u_2'' = 1$.

Encontramos que

$$u_7' = 8, u_1' + u_2' + \ldots + u_{10}' = 88,$$
$$u_7'' = 13, u_1'' + u_2'' + \ldots + u_{10}'' = 143.$$

Con un poco de suerte, podemos observar que

(*) $\quad u_1' + u_2' + \ldots + u_{10}' = 11u_7', \quad u_1'' + u_2'' + \ldots + u_{10}'' = 11u_7'',$

después intuir y, finalmente, probar que

(**) $\qquad u_1 + u_2 + \ldots + u_{10} = 11u_7.$

La prueba es ésta: verificamos directamente que

(***) $\qquad u_n = (u_2 - u_1)u_n' + u_1u_n''$

se mantiene para $n = 1$ y $n = 2$, y de aquí concluimos, utilizando la relación periódica, que vale también para $n = 3, 4, 5, \ldots 10$. Añadiendo las dos ecuaciones observadas (*) después de haber multiplicado la primera por $u_2 - u_1$ y la segunda por $u_1$, concluimos de (***) la (**) que deseábamos. La idea principal de la prueba: la solución general

$u_n$ de nuestra relación preriódica (más adecuadamente llamada una ecuación lineal con diferencia homogénea de segundo orden) es una combinación lineal de las dos soluciones particulares independientes $u_n'$ y $u_n''$ (como la integral general de una ecuación diferencial integral homogénea de segundo orden es una combinación lineal de dos integrales particulares).

**11.**
$$\int_0^\infty \frac{1}{1+x^\alpha}\,\frac{dx}{1+x^2} = \int_0^\infty \frac{x^{-\alpha}}{x^{-\alpha}+1}\,\frac{1}{x^{-1}+x}\,\frac{dx}{x}$$
$$= \int_0^\infty \frac{x^\alpha}{1+x^\alpha}\,\frac{dx}{1+x^2}$$
$$= \frac{1}{2}\int_0^\infty \frac{1+x^\alpha}{1+x^\alpha}\cdot\frac{dx}{1+x^2} = \frac{\pi}{4}$$

*independiente* de $\alpha$. Al pasar de la segunda forma a la tercera introducimos $x^{-1}$ como una nueva variable de integración. Para $\alpha = 0, \infty, -\infty$ la integral dada se reduce a

$$\int_0^\infty \frac{1}{2}\,\frac{dx}{1+x^2}, \qquad \int_0^1 \frac{dx}{1+x^2}, \qquad \int_1^\infty \frac{dx}{1+x^2},$$

respectivamente. Estos casos podrían sugerir la solución anterior.

**12.** El hecho más obvio de esta clase es:

$$\int_{-\infty}^{\infty} f(u)du = 0 \quad \text{si} \quad f(-u) = -f(u).$$

Establezcamos $u = \log x$, $f(\log x) = F(x)$:

$$\int_0^\infty F(x)x^{-1}dx = 0 \quad \text{si} \quad F(x^{-1}) = -F(x).$$

Esto sugiere la siguiente generalización:

$$\int_0^\infty g(x)\,[1+h(x)]x^{-1}dx = \int_0^\infty g(x)x^{-1}dx \text{ si } g(x^{-1}) = g(x),\ h(x^{-1}) = -h(x).$$

El ejemplo 11 es el caso particular:

$$g(x) = \frac{x}{2(1 + x^2)}, \qquad h(x) = \frac{1 - x^\alpha}{1 + x^\alpha}.$$

**13.** $0x = 0$, o $x^2 - 4 = (x - 2)(x + 2)$, etc.

**14.** $x = y = 8$; es suficiente probar $x = 8, 9, 10, 11$.

**15.** $x = y = z = w = 4$, por ensayos.

**17.** [Cf. Standford, 1948]. Los planos de simetría de un sólido regular pasan por su centro y dividen una esfera con el mismo centro en triángulos esféricos. Los tres radios que pasan por los tres vértices de este triángulo esférico pasan a través de un vértice, el centro de una cara y el punto medio de una arista, respectivamente. Los correspondientes ángulos del triángulo esférico son $\pi/v$, $\pi/f$ y $\pi/2$. Llamemos $c$ al lado (hipotenusa) del triángulo esférico opuesto al ángulo $\pi/2$. La razón del radio de la esfera inscrita al de la circunscrita es cos de $c$ y, por trigonometría esférica,

$$\cos c = \cot(\pi/f)\cot(\pi/v).$$

Los números $f$ y $v$ y el valor resultante de cos de $c$ se muestran en la siguiente tabla el Tetraedro, el Hexaedro (cubo), el Octaedro, el Dodecaedro y el Icosaedro.

| | T | H | O | D | I |
|---|---|---|---|---|---|
| $f$ = | 3 | 4 | 3 | 5 | 3 |
| $v$ = | 3 | 3 | 4 | 3 | 5 |
| $\cos c$ = | $\frac{1}{3}$ | $\frac{1}{\sqrt{3}}$ | | $\sqrt{\frac{5+2\sqrt{5}}{15}}$ | |

H. WEYL, *Symmetry*, Princeton, 1952, reproduce la figura original de Kepler; ver pág. 76, fig. 46.

**18.** Ver ejemplo 10.42.

**21.** (*a*) Llamamos a un determinante con $n$ filas simétrico-central si sus elementos $a_{j,k}$ satisfacen la condición

$$a_{j,k} = a_{n+1-j,\, n+1-k} \text{ para } j, k = 1, 2, 3, \ldots n.$$

*Un determinante simétrico-central con n filas es el producto de dos determinantes. Bien ambos factores tienen n/2 filas, bien un factor tiene* $(n + 1)/2$ *filas y el otro* $(n - 1)/2$ *filas, según sea n par o impar.* Ejemplos:

$$\begin{vmatrix} a & b \\ b & a \end{vmatrix} = (a + b)(a - b), \quad \begin{vmatrix} a & b & c \\ d & c & d \\ c & b & a \end{vmatrix} = \begin{vmatrix} a + c & b \\ 2d & c \end{vmatrix} (a - c),$$

$$\begin{vmatrix} a & b & c & d \\ e & f & g & h \\ h & g & f & e \\ d & c & b & a \end{vmatrix} = \begin{vmatrix} a + d & b + c \\ e + h & f + g \end{vmatrix} \quad \begin{vmatrix} f - g & e - h \\ b - c & a - d \end{vmatrix}.$$

Prueba: Poner $n = 2m$ o $n = 2m + 1$, según sea $n$ par o impar. Añadir la última columna a la primera, luego la penúltima a la segunda, y así sucesivamente, hasta que cambien las $m$ primeras columnas. Después de esto restar la primera fila de la última, luego la segunda fila de la penúltima, y así hasta que las últimas $m$ filas cambien. Estas operaciones presentan o un rectángulo $m \times (m + 1)$, o un cuadrado $m \times m$, consistente en elementos que desaparecen de la esquina inferior izquierda.

(*b*) La determinante de cuatro filas podría ser divisible por las determinantes de dos filas *sin* ser un producto, si estas determinantes de dos filas tuviesen un factor común. Con optimismo suponemos que no existe tal factor común: intentamos la suposición más sencilla y tenemos éxito.

**22.** Máximo optimismo: el coeficiente de cualquier valor de $h$ en el lado izquierdo es menor o igual que el coeficiente del mismo valor en el lado derecho. Esto es realmente el caso: después de la división por $4h^{1/2}$ el término constante es 1 a ambos lados y, para $n \geqq 1$, los coeficientes de $h^n$ son

$$\frac{3}{4}\,\frac{7}{8}\,\frac{11}{12} \cdots \frac{4n-1}{4n}\,\frac{1}{4n+1}, \qquad \frac{1}{4}\,\frac{5}{8}\,\frac{9}{4} \cdots \frac{4n-3}{4n}$$

a los lados respectivos. Obviamente,

$$3 \cdot 7 \cdot 11 \ldots (4n - 1) < 5 \cdot 9 \cdots (4n - 3)(4n + 1).$$

**23.** (*a*) Llamemos $P_n$ al valor aproximado obtenido con el método en cuestión cuando el cuadrado se subdivide en $n^2$ cuadrados más pequeños. Suponemos que $P_n$ puede ser ampliado a valores de $n^{-1}$:

$$P_n = Q_0 + Q_1 n^{-1} + Q_2 n^{-2} + \ldots$$

(«En general, una función puede ser ampliada a una serie de valores.» Cf. ejemplo 20.) A medida que $n \to \infty$, $P_n \to Q_0$ e inferimos que $Q_0 = Q$. Ahora bien, los cuatro puntos de la figura 11.6 están más próximos a una línea recta que los de la figura 11.5. Esta circunstancia *sugiere* que $Q_1 = 0$ y los términos $n^{-3}$, $n^{-4}$ son desdeñables incluso para $n$ pequeño. Esto lleva a

$$P_n \sim Q + Q_2 n^{-2}$$

que representa, si tomamos a $n^{-2}$ como abscisa y $P_n$ como ordenada, una línea recta (aproximadamente). En algunos casos más o menos semejantes se ha probado que el error de aproximación es del orden de $1/n^2$, y a la luz de tal analogía la intuición aparece menos tosca.

(*b*) Las columnas de la tabla siguiente contienen: (1) valores de $n$, (2) ordenadas, (3) diferencias de ordenadas, (4) abscisas, (5) diferencias de abscisas, (6) descensos calculados como la razón de (3) a (5), excepto que en (5) y (6) se ha omitido el signo —.

| (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|
| 2 | 0.0937 | | 0.2500 | | |
| | | 0.0248 | | 0.1389 | 0.1785 |
| 3 | 0.1185 | | 0.1111 | | |
| | | 0.0094 | | 0.0486 | 0.1934 |
| 4 | 0.1279 | | 0.0625 | | |
| | | 0.0045 | | 0.0225 | 0.2000 |
| 5 | 0.1324 | | 0.0400 | | |

(*c*) Es natural considerar $n = 5$ como el cálculo más seguro y $n = 4$ como el mejor de los que siguen. Si los puntos $(x_1, y_1)$ y $(x_2, y_2)$ están en línea recta con la ecuación $y = mx + b$, encontramos fácilmente (a partir de un sistema de dos ecuaciones para $m$ y $b$) que

$$b = \frac{y_1/x_1 - y_2/x_2}{1/x_1 - 1/x_2},$$

lo cual da por resultado, en el presente caso,

$$Q \sim \frac{25 \times 0.1324 - 16 \times 0.1279}{25 - 16} = 0.1404.$$

Si esperaba algo mejor que esto es usted demasiado vehemente.
Sin solución: **16, 19, 20.**

## CAPITULO XII

**1.** La semejanza con la tabla de la sección 3.12 (particiones del espacio) puede ser considerada más de cerca. Todas las tablas mencionadas, excepto la de la sección 3.1, que se refiere a los poliedros, están relacionadas con la evidencia inductiva que apoya una proposición $A$ de una determinada naturaleza: $A$ afirma que cierto enunciado $S_n$, cuyo significado depende de un entero $n$ variable, es verdadero para $n = 1, 2, 3, \ldots$ El examen inductivo de tal proposición $A$ procede, claro está, en cierto orden: primero probamos $S_1$, luego $S_2$, y así sucesivamente. Este orden es visible en la disposición de las tablas. Sin embargo, si examinamos una proposición referente a poliedros, como en la sección 3.1, no existe ese orden «natural». Podemos empezar nuestra investigación con el tetraedro, el cual, desde varios puntos de vista, puede ser considerado como el poliedro «más sencillo». Pero, ¿qué poliedro examinaremos seguidamente? No hay una razón convincente para considerar a algún poliedro como el «segundo en sencillez», ni a otro como el tercero, etc.

**2.** Si los casos 101 y 301 parecen llegar más allá de los ya verificados 1, 2, 3, ... 20 que los casos 21 y 22 (y en la investigación de Euler parece así), resulta razonable, y está en armonía con el patrón presentado en la sección 2, conceder más peso a la verificación de los casos 101 y 301 que a la de los casos 21 y 22.

**3.** (1) $c = p$: el triángulo degenera en una línea recta; $A = 0$.

(2) $c > p$ hace a $A$ imaginaria: no hay un triángulo con $c > p$.

(3) $a = b = c$: el triángulo es equilátero y $A^2 = 3a^4/16$, lo cual es correcto.

(4) $a^2 = b^2 + c^2$: el triángulo es rectángulo y

$$\begin{aligned} 16A^2 &= (a + b + c)(b + c - a)(a - b + c)(a + b - c) \\ &= [(b + c)^2 - a^2][a^2 - (b - c)^2] \\ &= (2bc)^2 \end{aligned}$$

o $A^2 = b^2c^2/4$, lo cual es correcto.

(5) $b = c = (h^2 + a^2/4)^{1/2}$: el triángulo es isósceles, con altura $h$, y

$$\begin{aligned} 16A^2 &= (a + 2b)(2b - a)a^2 \\ &= (4b^2 - a^2)a^2 \\ &= 4h^2a^2, \end{aligned}$$

que es correcto.

(6) La dimensión es correcta.

(7) La expresión para $A^2$ es simétrica en los tres lados $a$, $b$ y $c$, como debe ser.

**4.** (1) $d = 0$: el cuadrilátero se transforma en un triángulo y la fórmula enunciada se reduce a la de Herón; ejemplo 3.

(2) $d = p$: el cuadrilátero degenera en una línea recta; $A = 0$.

(3) $d > p$ hace a $A$ imaginaria: no existe un cuadrilátero con $d > p$.

(4) $a = b = c = d$ arroja $A^2 = a^4$, que es correcto para un cuadrado, pero incorrecto (demasiado grande) para un rombo: el cuadrado puede ser inscrito en un círculo, el rombo no.

(5) $c = a$, $d = b$ arroja $A^2 = (ab)^2$, que es correcto para un rectángulo, pero incorrecto (demasiado grande) para un paralelogramo oblicuo: el rectángulo puede ser inscrito en un círculo, el paralelogramo oblicuo no.

(6) En los casos anteriores, (4) y (5), la fórmula enunciada atribuye un valor demasiado grande al área de los cuadriláteros no inscribibles: esto concuerda con la sección 10.5 (2) y la 10.6 (3).

(7) La fórmula enunciada da la dimensión de $A$ correctamente.

(8) Según la fórmula propuesta, $A$ es simétrica en los cuatro lados $a$, $b$, $c$ y $d$: un cuadrilátero inscrito queda inscrito en el mismo círculo si los dos lados vecinos se intercambian. (Considerar los cuatro triángulos isósceles con vértice común en el centro del círculo cuyas bases son los cuatro lados.)

Estas observaciones no prueban la fórmula propuesta, naturalmente, pero la hacen muy plausible, de acuerdo con los patrones mostrados en la sección 2. Para una demostración ver el ejemplo 8.41.

**5.** (1) $a = b = c = e = f = g$: el tetraedro es regular; $V = 2^{1/2}\, a^3/12$.

(2) $e^2 = b^2 + c^2$, $f^2 = c^2 + a^2$, $g^2 = a^2 + b^2$: el tetraedro es «trirrectangular», es decir, las tres aristas $a$, $b$ y $c$ salen del mismo vértice y son perpendiculares entre sí; $V = abc/6$.

(3) $e = 0$, $b = c$, $f = g$: el tetraedro desaparece, se transforma en una figura plana, un triángulo; $V = 0$.

(4) $e = a$, $f = b$, $g^2 = c^2 = a^2 + b^2$: el tetraedro desaparece, se transforma en un rectángulo de lados $a$ y $b$; $V = 0$.

(5) Un caso particular más extendido que (4): el tetraedro se trans-

forma en un cuadrilátero, de lados $a$, $b$, $e$, $f$ y de diagonales $c$ y $g$. Luego, $V = 0$ da una relación entre los lados y las diagonales de un cuadrilátero general que también puede ser verificado directamente, aunque con menos facilidad.

(6) La dimensión es correcta.

(7) La expresión de $V$ *no* es simétrica en las seis aristas, pero no necesita serlo: las tres aristas $a$, $b$ y $c$ que salen del mismo vértice no juegan el mismo papel que las tres aristas $e$, $f$, $g$, que comprenden un triángulo (una cara). Podemos transformar, no obstante, la expresión propuesta en la siguiente:

$$\begin{aligned} 144V^2 = {} & a^2e^2\,(b^2 + f^2 + c^2 + g^2 - a^2 - e^2) \\ & + b^2f^2\,(c^2 + g^2 + a^2 + e^2 - b^2 - f^2) \\ & + c^2g^2\,(a^2 + e^2 + b^2 + f^2 - c^2 - g^2) \\ & - e^2f^2g^2 - e^2b^2c^2 - a^2f^2c^2 - a^2b^2g^2. \end{aligned}$$

Las tres primeras líneas corresponden a los tres pares de aristas opuestas; los cuatro términos de la última línea, a las cuatro caras del tetraedro: la nueva forma algebraica muestra toda la simetría (intercambiabilidad) de los datos debidos a la configuración geométrica. A propósito, la nueva forma (la corrección de la manipulación algebraica que la produce) puede ser también buscada en los casos particulares (1), (2), (3), (4).

**6.** (1) Encontramos para

$$s_1, \quad s_3, \quad s_5, \quad q, \quad r$$

los valores numéricos

$$6, \quad 36, \quad 276, \quad 11, \quad 6,$$

respectivamente, los cuales verifican ambas fórmulas.

(2) Más generalmente, hacemos

$$a + b + c = s_1 = p = 0.$$

Entonces

$$\begin{aligned} s_3 &= -3ab(a + b), \\ s_5 &= -5ab(a + b)\,(a^2 + ab + b^2), \\ q &= -a^2 - ab - b^2, \\ r &= -ab(a + b). \end{aligned}$$

y ambas fórmulas se verifican, puesto que

$$-\frac{3s_5}{5s_3} = -a^2 - ab - b^2, \qquad \frac{s_3}{3} = -ab(a+b).$$

(3) Más generalmente, hacemos a $c = 0$. Un cálculo más completo verifica ambas fórmulas.

(4) Más generalmente, hacemos a $b + c = 0$. Entonces, $s_1 = ,a$ $s_3 = a^3$, $s_5 = a^5$; denominador y numerador desaparecen en las expresiones propuestas para $q$ y $r$.

Para generalización y prueba ver *Journal des math. pures et appliquées*, serie 9, vol. 31 (1952), páginas 37-47.

**7.** El enunciado $B_4$ (con $r = h = 0$) está contenido como un caso particular en $B_2$ ($r = 0$) y en $B_3$ ($h = 0$). De aquí que si $B_2$ o $B_3$ son verdaderas, $B_4$ debe también ser verdadera. Por tanto, la verificación de $B_4$ depende de la de $B_2$ o $B_3$ y no nos aporta nueva información, ni, por consiguiente, nueva evidencia. Con todo, es deseable observar a $B_4$, pues redondea el cuadro.

**8.** Una atención cuidadosa a la derivación en los números 9-13 de la memoria de Euler, en la sección 6.2, muestra que $C_1^*$, $C_2$, ... $C_{20}^*$ se sigue de $C_1$, $C_2$, ... $C_{20}$ matemáticamente: por tanto, la verificación de $C_1^*$, $C_2^*$, ... $C_{20}^*$ no produce realmente nueva información o evidencia, pero la $C_{101}^*$ y $C_{103}^*$ sí.

**9.** El patrón concuerda esencialmente con el que introduciremos en el ejemplo 11; ver comentarios allí.

**10.** Si $8n + 3 = w^2 + 2p$, el entero $w$ es impar necesariamente. Por tanto, $w^2$ es de la forma $8n + 1$ y $p$ de la forma $4n + 1$. Euler probó que

$$p = u^2 + v^2;$$

de los dos enteros $u$ y $v$, uno debe ser impar y el otro par. De aquí que

$$2p = 2u^2 + 2v^2 = (u+v)^2 + (u-v)^2.$$

Ahora bien, $w$, $u + v$, y $u - v$ son impares. Hagamos

$$w = 2x - 1, u + v = 2y - 1, u - v = 2z - 1$$

y obtendremos

$$8n + 3 = (2x-1)^2 + (2y-1)^2 + (2z-1)^2$$

o

$$n = \frac{x^2 - x}{2} + \frac{y^2 - y}{2} + \frac{z^2 - z}{2}.$$

**13.** Sí, estaría en armonía con el patrón presentado en la sección 6. Sin solución: **11**, **12**, 14.

## CAPITULO XIII

**1.** El siguiente patrón es aplicable generalmente:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ falsa}\end{array}}{A \text{ falsa.}}$$

Aplicarlo sustituyendo no-$B$ por $B$. Se obtiene

$$\frac{\begin{array}{c}A \text{ implica no-}B \\ \text{no-}B \text{ falsa}\end{array}}{A \text{ falsa.}}$$

Hemos anotado en la sección 4 (5) la equivalencia:

«$A$ implica $B$» eq. «$A$ incompatible con no-$B$».

También esta equivalencia es aplicable generalmente. Aplicarla sustituyendo no-$B$ por $B$. Se obtiene

«$A$ implica no-$B$» eq. «$A$ incompatible con $B$».

Damos por sentado que la negación de no-$B$ es $B$ (puesto que la negación de $B$ es no-$B$) y, por tanto, sustituimos $B$ por no-(no-$B$). Sustituyendo $B$ por $A$ en la equivalencia anotada al final de la sección 4 (3) obtenemos también

«no-$B$ falsa» eq. «$B$ verdadera».

Sustituyendo los correspondientes enunciados por la dos premisas del último patrón, expresadas en el lado derecho de las dos equivalencias anteriores, obtenemos:

$$\frac{\begin{array}{c}A \text{ incompatible con } B \\ B \text{ verdadera}\end{array}}{A \text{ falsa.}}$$

Este es, en efecto, el patrón demostrativo de la sección 3.

**2.** Asumimos que el siguiente patrón es generalmente aplicable:

$$\frac{\begin{array}{c}A \text{ implica } B \\ B \text{ verdadera}\end{array}}{A \text{ más digna de crédito.}}$$

Sustituyamos $A$ por no-$A$ y $B$ por no-$B$. Así, obtenemos:

$$\frac{\begin{array}{c}\text{no-}A \text{ implica no-}B \\ \text{no-}B \text{ verdadera}\end{array}}{\text{no-}A \text{ más digna de crédito.}}$$

Recojamos las tres equivalencias siguientes:

«no-$A$ implica no-$B$» eq. «$B$ implica $A$»

«no-$B$ verdadera» eq. «$B$ falsa»

«no-$A$ más digna de crédito» eq. «$A$ menos digna de crédito».

La primera ha sido derivada en la sección 4 (5). La segunda fue mencionada en la sección 4 (3), bajo otra notación, con $A$ por $B$. La tercera fue establecida (sólo para el propósito presente) en la sección 5. Sustituyamos las premisas y la conclusión del último patrón considerado por los tres enunciados equivalentes acabados de expresar. Así, obtenemos:

$$\frac{\begin{array}{c}B \text{ implica } A \\ B \text{ falsa}\end{array}}{A \text{ menos digna de crédito.}}$$

Excepto un ligero cambio en las palabras (o notación) éste es el patrón heurístico presentado en la sección 2.

**3.** Empecemos por el mismo patrón que en el ejemplo 2. Sustituimos $B$ por no-$B$ (como en el ejemplo 1) y entonces obtenemos:

$$\frac{\begin{array}{c}A \text{ implica no-}B \\ \text{no-}B \text{ verdadera}\end{array}}{A \text{ más digna de crédito.}}$$

Recojamos las dos equivalencias siguientes:

«$A$ implica no-$B$» eq. «$A$ incompatible con $B$»

«no-$B$ verdadera» eq. «$B$ falsa».

La primera fue derivada en el ejemplo 1. La segunda se dio (con excepción de la notación) en la sección 4 (3). Sustituyamos las dos premisas del patrón considerado por los enunciados equivalentes acabados de manifestar, con lo que obtenemos:

$$\frac{\begin{array}{c}A \text{ incompatible con } B \\ B \text{ falsa}\end{array}}{A \text{ más digna de crédito.}}$$

Este es, en efecto, el patrón heurístico de la sección 3.

**4.** (*a*) En cualquier caso, $A$ implica $B$, donde el enunciado $B$ se define como sigue:

$B$. Las letras en que la palabra buscada de nueve letras se cruza con otras palabras del crucigrama hay que elegirlas entre las letras de la palabra TIREDNESS.

Interpretamos $B$ como limitada a dos lugares del diagrama propuesto (la última y la tercera antes de la última). Es conveniente distinguir dos casos.

Si consideramos las soluciones de las dos palabras cruzadas como definitivas encontramos que $B$ es verdadera y, por ello, verificamos una consecuencia de la conjetura $A$. Por tanto, consideramos $A$ como más digna de crédito de acuerdo con el patrón fundamental inductivo (sección 12.1).

Si, no obstante, consideramos las soluciones de las palabras cruzadas sólo como tentativas, $B$ únicamente se vuelve más digna de crédito. Por tanto, la versión matizada del patrón fundamental inductivo, definida en la sección 6, es más apropiada y, naturalmente, la evidencia de $A$ es más débil que en el primer caso.

(*b*) DISSENTER.

**5.** El patrón es el de la sección 13 (5):

El *factum probans* es realmente digno de crédito o comprensible bajo la asunción del *factum probandum.*

El *factum probans* es (mucho) menos digno de crédito o comprensible sin la asunción del *factum probandum.*

El *factum probans* mismo resulta probado.

---

Esto hace el *factum probandum* más digno de crédito.

La presentación que estamos considerando aquí parece ser más apropiada a un caso judicial, pero la presentación de la sección 10 es más conveniente para mostrar la conexión con la forma más usual de razonamiento inductivo en las ciencias físicas o en la investigación matemática.

**6.** Tenemos que tratar el cargo como una conjetura:

*A*. El pago del coche del oficial proviene del bolsillo del contratista.

Consideraremos como un hecho:

*B*. El retiro de la cuenta del contratista de una cantidad (875 dólares) igual al pago del coche del oficial, y que la fecha del retiro antecede en dos días a la fecha de dicho pago.

*B* con *A* es realmente mucho más comprensible que *B* sin *A*: si el retiro del dinero no estuviese conectado con el pago siguiente, la coincidencia exacta entre las dos cantidades y la cercana coincidencia de las fechas tiene que ser atribuida a mero azar. Tal azar no es imposible, pero improbable. La fuerza de la evidencia cae sobre este punto. El patrón de la sección 13 (5) parece adecuarse muy bien.

**7.** Llamemos a la señora White y al señor Black los «acusados». (No es que se le pueda imputar ningún cargo, pero, por lo menos, no son bien vistos por la señora Green.) La acusación de la señora Green, conducida por su sutil circunlocución, es, por supuesto:

*A*. Los acusados viven en doble adulterio.

Aceptamos como un hecho:

*B*. Los acusados tienen una larga conversación en la oscuridad sobre la valla.

Este hecho da, desgraciadamente, alguna evidencia circunstancial a *A* según razonable patrón (sección 13 (5)) si seguimos las dos premisas aceptadas:

*B* con *A* realmente más digna de crédito,

*B* sin *A* realmente menos digna de crédito.

Me temo que en el primer caso no se intenta en lo más mínimo conmover la fe de Suburbia. Sin embargo, algunos suburbanos pueden ver que, en la famosa tarde, los acusados estaban discutiendo el contrato en que ambos tenían un interés legítimo. Esto hace *B* tan digno de crédito sin *A* como con *A*, golpea la segunda premisa y echa fuera la alegada evidencia circunstancial. Este argumento puede ser que no afecte las murmuraciones de la señora Green, a pesar de que a mí me parece bastante razonable y típico. En todo caso, los abogados, al argüir contra una pieza de evidencia circunstancial, intentan a menudo rechazar la segunda premisa del patrón.

**8.** Consideremos otra conjetura:

*B*. Los acusados estaban familiarizados con la víctima tres años antes del delito.

Sería demasiado aventurado decir que $A$ está implicado en $B$, pero un enunciado más débil en esta dirección está obviamente justificado:

$A$ se hace más digna de crédito con $B$.

Ahora, $C$ no prueba $B$, pero, desde luego, hace a $B$ más digna de crédito.

A partir de las dos premisas manifestadas nos sentimos tentados de sacar esta conclusión: $A$ más digna de crédito.

Esto parece sugerir un nuevo patrón:

$$\frac{\begin{array}{c}A \text{ más digna de crédito con } B \\ B \text{ más digna de crédito}\end{array}}{A \text{ más digna de crédito.}}$$

La primera premisa de este patrón es más débil que la premisa correspondiente del patrón de la línea 2, columna (2) de la tabla I; la segunda premisa es la misma: la concluión debe ser también más débil. Ver, no obstante, el ejemplo 15.2.

Por cierto, el tamaño de la firma influye: si la firma es pequeña, $B$ se hace mucho más digna de crédito que en el caso contrario.

Sin solución: Ejemplos **9** al **20.**

## CAPITULO XIV

**1.** (*a*) Se sigue de $r_r + s_r = r_s + s_s = 1$.

(*b*) $r_r - r_s = s_s - s_r > 0$.

**2.** Con la notación del ejemplo 1, apropiadamente adaptado, $r_r - r_s = s_s - s_r < 0$.

**3.** $N\binom{n}{s} p^s q^{n-s}$, con $N = 26306$, $n = 12$, $p = 1/3$, $q = 2/3$ donde $s$ el número en la misma fila de la columna (1).

**4.** La misma expresión que en la solución del ejemplo 3, con los mismos valores numéricos para $N$ y $n$, y el mismo significado de $s$, pero $p = 0{,}3376986$, $q = 1 - p$.

**5.** (*a*) $N\mu^s e^{-\mu}/s!$ con $N = 30$, $\mu = 10$, y $s$ la entrada correspondiente en la columna (1).

(*b*) $- - - + + - - + - - - + + + - + - + + +$.

**6.** $6^{-3n}$.

**7.** El suceso compuesto que consiste en sacar seis puntos con cada uno de los tres dados cinco veces en sucesión ininterrumpida y tiene la probablidad $6^{-15}$ bajo la hipótesis de unos dados sin trampa: $n = 5$ en el ejemplo 6.

**8.** (*a*) $2\int_{\alpha}^{\infty} y\,dx = 1{,}983 \cdot 10^{-7}$,

donde $y = (2\pi)^{-1/2}\, e^{-x^2/2}$, $\alpha = 1377{,}5\,(pqn)^{-1/2}$,

$$p = 1/3, \qquad q = 2/3, \qquad n = 315672.$$

(*b*) $$\int_{\alpha}^{\beta} y\,dx = 1{,}506 \cdot 10^{-3}$$

con $\beta = -\alpha = 0{,}5\,(pqn)^{-1/2}$ e $y$, $p$, $q$ y $n$ tienen el mismo significado que en (*a*). Para calcular los valores numéricos tan exactamente como se dan aquí, las tablas más sencillas de la probabilidad integral no son suficientes.

**9.** Para encontrar artículos no defectuosos en la muestra: sólo uno defectuoso, precisamente dos defectuosos... precisamente $c$ defectuosos, respectivamente.

**10.** $d^2a/dp^2 = 0$ si la derivada del log $(da/dp)$ desaparece, lo cual produce la ecuación

$$\frac{c}{p} - \frac{n-1-c}{1-p} = 0,$$

y de aquí el valor dado al final de la sección 8 (1).

**11.** La probabilidad pedida es $10^{-n}$, *dado* que aceptemos una u otra de las siguientes suposiciones:

(I) Todas las secuencias posibles de $n$ figuras son igualmente probables. (Hay $10^n$ de tales secuencias.)

(II) Las distintas figuras de la secuencia son mutuamente independientes, y los diez casos posibles 0, 1, 2, ... 9 son igualmente probables para cada figura. (Aplicar repetidamente la regla de la sección 3 (5).)

Ambas suposiciones parece «naturales», pero esta suposición no es obligada lógicamente: la respuesta $10^{-n}$, aunque *fuertemente sugerida*, *no* está determinada matemáticamente.

**12.** Supongamos que $l$ letras diferentes pueden ser sacadas de una bolsa, con probabilidades $p_1$, $p_2$, ... $p_l$, respectivamente. Tenemos dos de estas bolsas y sacamos una letra de cada una: la probabilidad de una coincidencia es $p_1^2 + p_2^2 + \ldots + p_l^2$. En el caso de la hipótesis (II), $l = 17$ y $p_1, \ldots p_{17}$ pueden ser encontradas por cálculo real.

**13.** En ambos casos,

$$\sum_{k=n}^{10} \binom{10}{k} p^k q^{10-k} = 1 - \sum_{k=0}^{n-1} \binom{10}{k} p^k q^{10-k},$$

donde $q = 1 - p$; (*a*) $p = 0{,}0948$, (*b*) $p = 1/26 = 0{,}03846$.

**14.** En ambos casos, $np$ con $p = 0.0948$; (*a*) $n = 450$, (*b*) $n = 90$.

**15.** $(npq)^{1/2}$ para $n = 90$, $p = 0{,}0948$, $q = 1 - p$. El cálculo de la desviación normal 7,60 está basado en una fórmula que no se encuentra en los libros de texto. Con la anotación del ejemplo 12 establecemos

$$p = p_1^2 + p_2^2 + \ldots + p_l^2,$$

$$p' = p_1^3 + p_2^3 + \ldots + p_l^3,$$

$$\sigma^2 = n(n-1)\, w[p(1-p) + 2(n-2)\,(p' - p^2)]/2.$$

Luego, $n = w = 10$, $p = 0{,}0948$, $p' = 0{,}01165$ arroja $\sigma = 7{,}60$.

**16.** Suponemos que los sesenta ensayos con la moneda son independientes y aplican la regla presentada en la sección 3 (5) repetidamente.

**17.** Generalizar la tabla numérica propuesta

| | | |
|---|---|---|
| $s$ | $r$ | $n$ |
| $s'$ | $r'$ | $n'$ |
| $S$ | $R$ | $N$ |

e interpretarla como sigue. Hay $N = R + S = n + n'$ cartas, entre las cuales $R = r + r'$ cartas son rojas y $S = s + s'$ son negras. Las cartas se distribuyen al azar entre dos jugadores; uno recibe $n$ cartas, y el otro $n'$ cartas. ¿Cuál es la probabilidad de que el primer jugador reciba $r$ cartas rojas y $s$ cartas negras, y el otro jugador reciba $r'$ cartas

rojas y $s'$ negras? (Naturalmente, $r + s = n$, $r' + s' = n'$.) La respuesta es, como sabemos,

$$\frac{\binom{R}{r}\binom{S}{s}}{\binom{N}{n}} = \binom{n}{r} \frac{R(R-1)\ldots(R-r+1)\,S(S-1)\ldots(S-s+1}{N(N-1)\ldots(N-n+1)}$$

$$= \frac{1}{N!}\,\frac{R!\;S!\;n!\;n'!}{r!\;s!\;r'!\;s'!}\,.$$

Sólo cuatro cantidades fuera de las nueve contenidas en la tabla pueden ser dadas arbitrariamente; los valores de las cinco restantes se siguen de las relaciones escritas anteriormente. Tomamos los números $n = 9$, $n' = 11$ y $S = 8$ como dados (el número de pacientes que recibe cada tratamiento y el número total de casos fatales) de los cuales se siguen $N = 20$ y $R = 12$. No obstante, tomamos $s' = 2, 1, 0$, sucesivamente (número de desgracias con el segundo tratamiento $\leqq 2$). A partir de la fórmula calculamos la probabilidad en cada uno de los tres casos (no más de 2 fatalidades con el nuevo tratamiento, 2, 1 o 0 cartas negras para el señor Newman) y añadiendo las probabilidades para estos acontecimientos mutuamente excluyentes encontramos:

$$\frac{12!\,8!\,9!\,11!}{20!}\left[\frac{1}{3!\;6!\;9!\;2!} + \frac{1}{2!\;7!\;10!\;1!} + \frac{1}{1!\;8!\;11!\;0!}\right]$$

$$= \left[\binom{12}{9}\binom{8}{2} + \binom{12}{10}\binom{8}{1} + \binom{12}{11}\binom{8}{0}\right] \Big/ \binom{20}{11} = \frac{335}{8398}\,.$$

**18.** Por una ampliación obvia del razonamiento de la sección 3 (5) (tridimensional análogo a la figura 14.2) el número de casos posibles es $n^3$. Todos los casos favorables, es decir, todas las soluciones admisibles de la ecuación $Z = X + Y$ pueden ser enumeradas como sigue:

$$\begin{aligned}
2 &= 1 + 1 \\
3 &= 1 + 2 = 2 + 1 \\
4 &= 1 + 3 = 2 + 2 = 3 + 1 \\
&\;\;\cdot \qquad \cdot \qquad \cdot \qquad \cdot \\
n &= 1 + (n-1) = 2 + (n-2) = \ldots = (n-1) + 1.
\end{aligned}$$

De aquí que el número de casos favorables sea

$$1 + 2 + 3 + \ldots + (n-1) = n(n-1)/2$$

y la probabilidad pedida

$$\frac{n(n-1)/2}{n^3} = \frac{n-1}{2n^2}.$$

**21.** La probabilidad de que una muestra de 38 de una población infinita contenga 30 o más defectivos es

$$\sum_{i=30}^{38} \binom{38}{i} p^i (1-p)^{38-i} < \binom{38}{8} p^{30}(1-p)^8 \sum_{i=0}^{\infty} \left(\frac{p}{1-p}\right)^i$$

$$= \binom{38}{8} \frac{(1-p)^9}{1-2p} p^{30}$$

$$\sim 4{,}61 \times 10^{-23}$$

previsto que el 100 por $100p = 1$ por 100 de la población sea defectivo. La probabilidad estimada es la verosimilitud de la aserción diaria a la luz de la observación oficial, calculada bajo las suposiciones más sencillas.

**22.** La probabilidad pedida es

$$4\left\{\left(\frac{2}{12 \times 60}\right)^3 + 3\left(\frac{2}{12 \times 60}\right)^2 \left(1 - \frac{2}{12 \times 60}\right)\right\} \sim 0{,}0000924.$$

(Cualquiera de los cuatro relojes podría ser el que, entre los tres que van aproximadamente igual, se muestra más atrasado. La posibilidad de que los cuatro relojes muestren horas con menos de dos minutos de diferencia cuenta en el primer término del paréntesis.)

**23.** El entero $a$ toma uno de los valores

$$-n, \ldots, -2, -1, 0, 1, 2, \ldots n.$$

Suponemos que estos $2n + 1$ valores son igualmente probables, hacemos la correspondiente suposición para $b$, $c$, $d$, $e$ y $f$, y asumimos también que $a$, $b$, $c$, $d$, $e$ y $f$ son mutuamente independientes. Sólo ahora, después de haber dado un significado preciso a «elegir al azar», podemos proceder a resolver el problema.

Hay sólo una solución si, y solamente si, $ad - bc \neq 0$. Podemos desdeñar $e$ y $f$: hay $(2n + 1)^4$ casos posibles. Contamos los casos *desfavorables* distinguiendo dos posibilidades.

(1) $a = 0$. Luego $bc = 0$, $d$ es arbitrario y hay $(2n + 2n + 1)(2n + 1)$ casos.

(II) $a \neq 0$. Luego $a$ puede tomar $2n$ valores, $b$ y $c$ son arbitrarios y $d$ está únicamente determinado por $a$, $b$ y $c$: hay $2n(2n+1)^2$ casos.

La probabilidad pedida es

$$1 - \frac{(4n+1)(2n+1) + 2n(2n+1)^2}{(2n+1)^4}$$

$$= 1 - \frac{4n^2 + 6n + 1}{8n^3 + 12n^2 + (6n+1)} > 1 - \frac{1}{2n}$$

y, así, tiende a 1 como $n$ tiende a $\infty$. Esto da otro significado preciso al enunciado: «un sistema de dos ecuaciones con dos incógnitas tiene, *en general*, sólo una solución». Cf. sección 11.3, ejemplo 11.16.

**24.** La verificación de la O da más confianza en TOWER que la verificación de la E. Puesto que O es menos frecuente que E, aquélla es interpretada menos fácilmente en el crucigrama como coincidencia azarosa que E.

**25.** Tabulando las diferencias entre los números sucesivos de cada columna, y luego otra vez las diferencias de las diferencias (las llamadas «segundas diferencias»), obtenemos:

| I | | | II | | |
|---|---|---|---|---|---|
| 1005 | | | 1004 | | |
| | 28 | | | 34 | |
| 1033 | + | 14 | 1038 | | 0 |
| | 42 | | | 34 | |
| 1075 | — | 11 | 1072 | | 0 |
| | 31 | | | 34 | |
| 1106 | — | 5 | 1106 | — | 1 |
| | 26 | | | 33 | |
| 1132 | + | 21 | 1139 | + | 1 |
| | 47 | | | 34 | |
| 1179 | — | 21 | 1173 | — | 1 |
| | 26 | | | 33 | |
| 1205 | | 0 | 1206 | | 0 |
| | 26 | | | 33 | |
| 1231 | + | 17 | 1239 | — | 1 |
| | 43 | | | 32 | |
| 1274 | — | 16 | 1271 | | 0 |
| | 27 | | | 32 | |
| 1301 | | | 1303 | | |

Las primeras diferencias muestran bastante claramente que II es regular, aunque no así I. Sin embargo, las segundas diferencias son todavía más sugestivas: II muestra un mínimo de irregularidad debida a los inevitables errores de redondeo, mientras las segundas diferencias de I varían de signo y son bastante grandes. Tal «diferenciación» es una importante operación de comprobación en la construcción de tablas numéricas. He aquí dos observaciones.

(1) En la tabla de una función $f(x)$ las primeras diferencias están conectadas con $f'(x)$, y las segundas diferencias con $f''(x)$, por el teorema del valor medio. Esto nos da una oportunidad para comprobar las diferencias.

(2) Los últimos decimales de I son, en efecto, los primeros diez decimales de $\pi$ en orden inverso. La consideración de que los decimales sucesivos de $\pi$ se conducen *como si* se produjeran por azar se ha expresado muchas veces en distintas formas.

**27.** Asumimos que $A$ es independiente de $B$:

$$\Pr\{A/B\} = \Pr\{A/\bar{B}\}. \tag{6}$$

(Empezamos con (6) para evitar interferencias con la numeración del ejemplo 26.)

Utilizando (4), (2), (6) y (3) (por este orden) obtenemos

$$\begin{aligned} \Pr\{A\} &= \Pr\{AB\} + \Pr\{A\bar{B}\} \\ &= \Pr\{B\}\Pr\{A/B\} + \Pr\{\bar{B}\}\Pr\{A/\bar{B}\} \\ &= \Pr\{A/B\}(\Pr\{B\} = \Pr\{\bar{B}\}) \\ &= \Pr\{A/B\}. \end{aligned} \tag{7}$$

De (2), (7) y $\Pr\{A\} \neq 0$ se sigue que

$$\begin{aligned} \Pr\{A\}\Pr\{B/A\} &= \Pr\{B\}\Pr\{A\} \\ \Pr\{B/A\} &= \Pr\{B\}. \end{aligned} \tag{8}$$

Utilizando de nuevo (4) y (2) como en (7), y utilizando también (8), (3) y $\Pr\{\bar{A}\} \neq 0$, encontramos que

$$\begin{aligned} \Pr\{B\} &= \Pr\{A\}\Pr\{B/A\} + \Pr\{\bar{A}\}\Pr\{B/\bar{A}\} \\ (1 - \Pr\{A\})\Pr\{B\} &= \Pr\{\bar{A}\}\Pr\{B/\bar{A}\} \\ \Pr\{B\} &= \Pr\{B/\bar{A}\}; \end{aligned} \tag{9}$$

(6), (7), (8) y (9) muestran la conclusión pedida.

**28.** Si $A$ y $B$ son mutuamente independientes,

$$\Pr\{AB\} = \Pr\{A\}\Pr\{B\}.$$

Esto se sigue de la regla (2) del ejemplo 26 y la definición (II) del ejemplo 27.

**29.** (*a*) $\Pr\{A\}$, $\Pr\{A/B\}$, $\Pr\{A/\bar{B}\}$, $\Pr\{B\}$, $\Pr\{B/A\}$, $\Pr\{B/\bar{A}\}$,

(I) $\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$,

(II) $\frac{1}{3}$, $1$, $\frac{1}{5}$, $\frac{1}{6}$, $\frac{1}{2}$, $0$,

(*b*) $$\Pr\{AB\} = \Pr\{A\}\Pr\{B/A\} = \Pr\{B\}\Pr\{A/B\},$$

(I) $$\tfrac{1}{3} \cdot \tfrac{1}{2} = \tfrac{1}{3} \cdot \tfrac{1}{3},$$

(II) $$\tfrac{1}{3} \cdot \tfrac{1}{2} = \tfrac{1}{6} \cdot 1,$$

(*c*) Las fórmulas se siguen generalmente del ejemplo 26 (4), (2). Numéricamente.

$$\Pr\{A\} = \Pr\{B\}\Pr\{A/B\} + \Pr\{\bar{B}\}\Pr\{A/\bar{B}\},$$

(I) $$\tfrac{1}{3} = \tfrac{1}{2} \cdot \tfrac{1}{3} + \tfrac{1}{2} \cdot \tfrac{1}{3},$$

(II) $$\tfrac{1}{3} = \tfrac{1}{6} \cdot 1 + \tfrac{5}{6} \cdot \tfrac{1}{5}.$$

$$\Pr\{B\} = \Pr\{A\}\Pr\{B/A\} + \Pr\{\bar{A}\}\Pr\{B/\bar{A}\},$$

(I) $$\tfrac{1}{2} = \tfrac{1}{3} \cdot \tfrac{1}{2} + \tfrac{2}{3} \cdot \tfrac{1}{2},$$

(II) $$\tfrac{1}{6} = \tfrac{1}{3} \cdot \tfrac{1}{2} + \tfrac{2}{3} \cdot 0.$$

(*d*) Con (I), $A$ y $B$ son mutuamente independientes, con (II) no lo son.

Sin solución: **19, 20, 26, 30, 31, 32, 33.**

## CAPITULO XV

**1.** Puesto que $H$ implica $A$ y $B$,

$$\Pr\{A/H\} = 1, \qquad \Pr\{B/H\} = 1.$$

Por tanto, por la regla (2) del ejemplo 14.26,

$$\Pr\{H\} = \Pr\{A\}\Pr\{H/A\}, \qquad \Pr\{H\} = \Pr\{B\}\Pr\{H/B\}.$$

Eliminando $\Pr\{H\}$ obtenemos

$$\Pr\{A\}\Pr\{H/A\} = \Pr\{B\}\Pr\{H/B\}.$$

Si consideramos la relación entre $H$ y $A$, y también la relación entre $H$ y $B$, como invariables, y, por tanto, $\Pr\{H/A\}$ y $\Pr\{H/B\}$ como constantes, si $\Pr\{B\}$ aumenta, $\Pr\{A\}$ también aumentará por virtud de nuestra última ecuación.

**2.** Por las fórmulas del ejemplo 14.26 (2), (3), (4)

$$\Pr\{A\} = \Pr\{A/\bar{B}\} + \Pr\{B\}(\Pr\{A/B\} - \Pr\{A/\bar{B}\}).$$

Tomamos $\Pr\{A/B\}$ y $\Pr\{A/\bar{B}\}$ como dado, pero $\Pr\{\bar{B}\}$ aumenta. (Esto puede verse con bastante naturalidad en el caso considerado en el ejemplo 13.8.) El aumento de $\Pr\{B\}$ implica un aumento correspondiente de $\Pr\{A\}$ *si*

$$(*) \qquad \Pr\{A/B\} > \Pr\{A/\bar{B}\}.$$

En el caso concreto del ejemplo 13.8 esta desigualdad parece aceptable. Sin embargo, el enunciado del patrón dado en la solución del ejemplo 13.8 parece inaceptable. Aquí tenemos un enunciado que parece mejor y está en línea con la fórmula:

$A$ más digno de crédito con $B$ que sin $B$
$B$ (se hace) más digno de crédito
---
$A$ (se hace) más digno de crédito.

Si $A$ es una consecuencia de $B$, $\Pr\{A/B\} = 1$, (*) es correcto, y el patrón de la sección 13.7, tabla I, línea 2, columna 2.

**3.** Revisar los pocos casos en que hemos representado patrones de razonamiento plausible mediante fórmulas del cálculo de probabilidad: secciones 6-10 y los ejemplos 1-2. La monotonicidad y la continuidad afirmada por el ejemplo 13.10 es obvia en los casos que se siguen del siguiente hecho matemático: Si $a$, $b$, $c$, $d$, $x_1$, $x_2$ son constantes reales, $ad - bc \neq 0$, y la función $y$ de $x$, definidas por

$$y = \frac{ax + b}{cx + d},$$

tiene la propiedad de que $0 \leqq y \leqq 1$ para $x_1 < x < x_2$, luego $y$ es estrictamente monotónica y continua para $x_1 \leqq x \leqq x_2$. Con respecto a la generalidad de las observaciones del ejemplo 13.10 observamos que:

si $y$ no es una función fraccional lineal, sino, más generalmente, una función racional de $x$, es todavía necesariamente una función continua de $x$, excepto en el punto en que se hace infinito, aunque no necesariamente una función monotónica.

Sin solución: **4, 5, 6, 7, 8, 9.**

## CAPITULO XVI

**2.** [Stanford, 1951.] El cuadrilátero tiene que ser convexo. Llamemos I, II, III, IV, los triángulos en que es dividido por sus diagonales; (I), (II), (III), (IV), las áreas de los cuatro triángulos, respectivamente, y $p, q, r, s$, la longitud de las cuatro líneas rectas trazadas desde la intersección de las diagonales a los cuatro vértices del cuadrilátero. Nombrar y enumerar en «orden cíclico» de modo que el lado de longitud $p$ sea común a IV y I, $q$ a I y II, $r$ a II y III, $s$ a III y IV; I es opuesto a III, II a IV, $p + r$ es la longitud de una diagonal, $q + s$ la de la otra. $p$ y $q$ comprenden el ángulo $\alpha$. Luego,

$$2\,(\mathrm{I}) = pq \operatorname{sen} \alpha, \qquad 2(\mathrm{II}) = qr \operatorname{sen} \alpha,$$

$$2\,(\mathrm{III}) = rs \operatorname{sen} \alpha, \qquad 2(\mathrm{IV}) = sp \operatorname{sen} \alpha.$$

De aquí,

($a$) (I) (III) = (II) (IV).

($b$) La base de I es paralela a la de III si, y sólo si,

$$p/q = r/s \quad \text{o} \quad (\mathrm{II}) = (\mathrm{IV}).$$

($c$) El cuadrilátero es un paralelogramo si, y sólo si,

$$p = r, \quad q = s \quad \text{o} \quad (\mathrm{I}) = (\mathrm{II}) = (\mathrm{III}) = (\mathrm{IV}).$$

**3.** [Stanford, 1949.] ($a$) Sea $a$ un lado del triángulo equilátero. Uniendo el punto interior a sus tres vértices lo dividimos en tres triángulos con áreas que, sumadas, dan el área total $ax/2 + ay/2 + az/2 = ah/2$. Dividida por $a/2$. Ver la figura 8.8.

($b$) Un punto interior dentro de un tetraedro regular de altura $h$ tiene las distancias $x, y, z$ y $w$ desde las cuatro caras, respectivamente. Luego, $x + y + z + w = h$. La prueba es análoga: divide el tetraedro regular en cuatro tetraedros.

($c$) La relación permanece válida en los casos ($a$) y ($b$) para puntos exteriores, teniendo en cuenta que las distancias $x, y, z$ (y $w$) se toman

con el signo propio: + cuando un espectador colocado en el punto ve el lado (cara) desde dentro, — cuando lo ve desde fuera. La prueba es, esencialmente, la misma.

**4.** [Stanford, 1946.] (*a*) (I) y (IV) son generalmente verdaderos, pero (II) y (III) no lo son; ver (*b*).

(*b*) (II) y (III) son falsos: el rectángulo y el rombo son contraejemplos, respectivamente.

(*c*) (II) y (III) son verdaderos para pentágonos.

(*d*) (II) y (III) son verdaderos para polígonos con un número impar de lados, tres o cinco o siete..., como se sigue de (II') y (III').

(II') Si un polígono inscrito en un círculo es equiangular, dos lados cualesquiera separados por un solo lado interpuesto son iguales. Por tanto, si el número de lados es par, igual a $2m$, o todos los $2m$ lados son iguales, o $m$ lados son iguales a $a$ y los $m$ restantes a $b$, $a \neq b$, y ningunos dos lados que tengan un vértice común son iguales.

(III') Si un polígono circunscrito a un círculo es equilátero, cualesquiera dos ángulos separados por un solo lado interpuesto son iguales. Por tanto, si el número de ángulos es par, igual a $2m$, todos los $2m$ ángulos son iguales, o $m$ ángulos son iguales a $\alpha$ y los $m$ ángulos restantes iguales a $\beta$, $\alpha \neq \beta$, y no hay dos ángulos que teniendo un lado común sean iguales.

Para probar (I), (II'), (III'), (IV), unir el centro del círculo a dos vértices del polígono, trazar perpendiculares desde el centro a los lados y sacar los triángulos congruentes.

**5.** [Stanford, 1947.] La relación entre (*a*) y (*b*) parece similar a la que hay entre (*b*) y (*c*); la similaridad de la izquierda parece más clara que la de la derecha. De aquí que sea natural buscar una transición de (*a*) a (*b*) que podría servir de transición de (*b*) a (*c*). He aquí tal transición: si

$$\alpha + \beta + \gamma = \pi$$

$$2\alpha + 2\beta + 2\gamma = 2\pi \neq \pi, \text{ pero}$$

$$(\pi - 2\alpha) + (\pi - 2\beta) + (\pi - 2\gamma) = \pi.$$

Dando por sentado que (*a*) es válido para cualesquiera tres ángulos $\alpha, \beta, \gamma$, que suman $\pi$, obtenemos (*b*) sustituyendo estos ángulos $\pi - 2\alpha$, $\pi - 2\beta, \pi - 2\gamma$, respectivamente, y pasamos de (*b*) a (*c*) utilizando la misma sustitución.

Queda por verificar (*a*), lo que puede ser hecho de muchas maneras,

por ejemplo, como sigue. Ponemos $2u$, $2v$ y $\pi - 2u - 2v$ en el lugar de $\alpha$, $\beta$ y $\lambda$, respectivamente; transformamos $(a)$ en

$$\operatorname{sen} u \cos u + \operatorname{sen} v \cos v = [2 \cos u \cos v - \cos (u + v)] \operatorname{sen} (u + v).$$

Emplear los teoremas de suma de senos y cosenos.

**6.** [Stanford, 1952]. De acuerdo a los cuatro propósitos, el volumen de los truncados sería

$$\text{(I)} \quad [(a + b)/2]^2 h$$

$$\text{(II)} \quad [(a^2 + b^2)/2]h$$

$$\text{(III)} \quad [a^2 + b^2 + (a + b)^2/4]h/3$$

$$\text{(IV)} \quad [a^2 + b^2 + ab]h/3,$$

respectivamente. Si $b = a$, se alcanza un prisma truncado de volumen $a^2h$: (I), (II), (III), (IV), están de acuerdo en dar el resultado correcto. Si $b = 0$ se alcanza una pirámide truncada de volumen $a^2h/3$: sólo (IV) da esto y, así, (I), (II), (III), que dan valores diferentes deben ser incorrectos. Que (IV) es generalmente correcto debe ser aún probado; ver los libros de texto.

**10.** (1) Haciendo la suma con papel y lápiz de la manera más usual, empezamos desde arriba de la última columna y procedemos hacia abajo; en el ejemplo dado los dos primeros pasos son $3 + 7 = 10$, $10 + 0 = 10$. Luego procedemos hacia abajo en la próxima columna, y así sucesivamente. (2) Tomamos las columnas en el mismo orden que antes, pero empezamos por abajo y procedemos hacia arriba en cada una. (3) Tomamos las columnas de nuevo en el mismo orden, pero hacemos cada columna dos veces, una hacia abajo y otra hacia arriba; observamos el resultado cuando vamos hacia abajo y luego lo comprobamos cuando vamos hacia arriba. (4) Sumamos primero el 1, 3, 5, ... y luego el 2, 4, 6, ... y, finalmente, las dos sumas obtenidas. (5) Hacer la suma, una vez a mano y otra con máquina. Y así sucesivamente.

**11.** 59.

**12.** $10^{-9}$. (No $10^{-3}$ o $10^{-7}$.)

**Sin solución: 1, 7, 8, 9, 13.**

# Bibliografía

## I. CLASICOS

EUCLIDES, *Elements*. La edición abreviada de la Everyman's Library es aquí suficiente. «EUCLIDES, III, 7» se refiere a la proposición 7 del libro III de los *Elementos*.
DESCARTES, *Oeuvres*, editadas por Charles Adam y Paul Tannery. La obra *Regulae ad Directionem Ingenii*, vol. 10, págs. 359-469, es de interés especial.
EULER, *Opera Omnia*, editada por la «Societas scientiarum naturalium Helvetica».
LAPLACE, *Oeuvres complètes*. La «introducción» del vol. 7, págs. V-CLIII, también impresa por separado (y mejor conocida) bajo el título *Essai philosophique sur les probabilités*, es de especial interés.

## II. ALGUNOS LIBROS DE TENDENCIA SIMILAR

R. COURANT y H. ROBBINS, *What is mathematics?*
H. RADEMACHER y O. TOEPLITZ, *Von Zahlen und Figuren.*
O. TOEPLITZ, *Die Entwicklung der Infinetisimalrechnung*; de interés especial.

## III. OBRAS DEL AUTOR RELACIONADAS

*Libros*:

1. *Aufgaben und Lehrsätze aus der Analysis*, 2 vols., Berlín, 1925. Juntamente con G. Szegö. Reimpresa en Nueva York en 1945.
2. *How to Solve It*, Princeton, 1945. La quinta edición está ligeramente alargada.
3. Wahrscheinlichkeitsrechnung, Fehlerausgleichung, Statistik. Del *Abderhalden's Handbuch der biologischen Arbeitsmethoden*, Abt. V, Teil 2, págs. 669-758.

*Artículos*:

1. «Geometrische Darstellung einer Gedankenkette». *Schweizerische Pädagogische Zeitschrift*, 1919, 11 págs.
2. «Wie sucht man die Lösung mathematischer Aufgaben?». *Zeitschrift fur mathematischen und naturwissenschaftlichzn Unterricht*, vol. 63, 1932, págs. 159-169.
3. «Wie sucht man die Losung mathematischer Aufgaben». *Acta Psychologica*, volúmen 4, 1938, págs. 113-170.
4. «Heuristic reasoning and the theory of probability». *American Mathematical Monthly*, vol. 48, 1941, págs. 450-465.
5. On Patterns of Plausible Inference». *Courant Anniversary Volume*, 1948, páginas 277-288.
6. «Generalization, Specialization, Analogy». *American Mathematical Monthly*, volúmen. 55, 1948, págs., 241-243.

7. «Preliminary remarks on a logic of plausible inference». *Dialectica*, vol. 3, 1949, páginas 28-35.

8. «With, or without, motivation?». *American Mathematical Monthly*, vol. 56, 1949, págs., 684-691.

9. «Let us teach guessing». *Etudes de Philosophie des Sciences, en hommage à Ferdinand Gonseth*, 1950, págs. 147-154. Editions du Griffon, Neuchatel, Suiza.

10. «On plausible reasoning». *Proceedings of the International Congress of Mathematicians*, 1950, vol. 1, págs. 739-747.

## IV. PROBLEMAS

Entre los ejemplos propuestos para solución hay algunos cogidos de la *William Lowell Putnam Mathematical Competition* o la *Stanford University Competitive Examination in Mathematics*. Este hecho se indica al principio de la solución con el año en que el problema se propuso, como «Putnam, 1948» o «Stanford, 1946». Los problemas de la Putnam Examination se publican anualmente en la *American Mathematical Monthly* y la mayoría de los de la Stanford son también publicados allí.

mos, etc., con respectivas ilustraciones. En efecto, el autor, más que gastar el tiempo arguyendo en pro de su tesis, se limita a presentar ejemplos con los que ayudar a los estudiantes a *aprender intuyendo* y a los profesores a *enseñar intuyendo* también.

La segunda parte del libro presenta los *patrones* o «reglas» del razonamiento plausible, basadas sobre la probabilidad matemática. Ésta nos da siempre una tosca indicación del grado de confianza de las conclusiones a que hemos llegado por otros métodos de razonamiento plausible, aunque como él mismo apunta, sea imposible obtener una cifra exacta de la credibilidad de una proposición.

Los ejemplos y problemas propuestos por el autor están llenos de interés, tratando siempre de desafiar la curiosidad, el juicio y las facultades creadores del lector de un modo que no es frecuente en los libros de texto sobre matemáticas. La presentación del libro es estimulante, hecha en un estilo sencillo, flexible y, en muchas ocasiones, lleno de humor.

*

GEORGE POLYA nació en Budapest (Hungría) en 1887. En 1912 se doctoró en Matemáticas en la Universidad de Budapest; estudió también en las Universidades de Viena, Gotinga y París. Ha vivido mucho tiempo en Zurich (Suiza), donde ha enseñado en diversas instituciones. En 1940 se trasladó a Estados Unidos, donde, desde 1942, enseñó en la Universidad de Stanford (California); al retirarse en 1953 fue nombrado *profesor emeritus* de dicha Universidad.